VOLUME 22 NUMBER 3 SEPTEMBER 1959

SOCIOMETRY

A JOURNAL OF RESEARCH IN SOCIAL PSYCHOLOGY

Articles by

ROSEN and D'ANDRADE
SCOTT
STEINER
NEFZGER and LILIENFELD
SOMMER
GOODCHILDS
SIROTA

Published by

THE AMERICAN SOCIOLOGICAL SOCIETY

SOCIOMETRY

EDITOR

LEONARD S. COTTRELL, JR.

EDITOR-ELECT

JOHN A. CLAUSEN

ASSOCIATE EDITORS

EDGAR F. BORGATTA
LAUNOR F. CARTER
LEON FESTINGER
NELSON N. FOOTE
NATHANIEL L. GAGE
ERVING GOFFMAN
FREDERICK MOSTELLER
HERBERT A. SIMON
FRED L. STRODTBECK
GUY E. SWANSON

EDITORIAL CONSULTANTS

WENDELL BELL
PETER M. BLAU
ORVILLE G. BRIM, JR.
ROGER W. BROWN
JAMES S. COLEMAN
SANFORD M. DORNBUSCH
FRED E. FIEDLER
ALLAN R. HOLMBERG
ALEX INKELES
MELVIN L. KOHN
WILLIAM W. LAMBERT
DUNCAN MACRAE, JR.
JAMES G. MARCH
THEODORE R. SARBIN
STANLEY SCHACHTER
WILLIAM A. SCOTT
SIDNEY SIEGEL
SHIRLEY A. STAR
RALPH H. TURNER
EUGENE A. WEINSTEIN
ROBIN M. WILLIAMS, JR.
ROBERT F. WINCH
MARIAN RADKE YARROW

EXECUTIVE OFFICER

MATILDA WHITE RILEY

Published four times a year

March, June, September, December

by

THE AMERICAN SOCIOLOGICAL SOCIETY

at The Boyd Printing Co., Inc., 49 Sheridan Ave., Albany 10, N. Y.

Sociometry was founded in 1937 by J. L. Moreno

Communications for the Editors and all manuscripts should be addressed to John A. Clausen, The National Institutes of Health, The Clinical Center, Room 2N208, Bethesda 14, Maryland. Business communications should be addressed to Matilda White Riley, The American Sociological Society, New York University, Washington Square, New York 3, New York.

Copyright © 1959 by THE AMERICAN SOCIOLOGICAL SOCIETY

Second class postage paid at Albany, N. Y.

SOCIOMETRY

A JOURNAL OF RESEARCH IN SOCIAL PSYCHOLOGY

VOLUME 22 · NUMBER 3 · SEPTEMBER 1959

CONTENTS

SOCIOMETRY

EDITORIAL POLICY

Sociometry is concerned with the entire range of interests and problems represented by research in social psychology. It is the policy of the editors to seek those manuscripts for publication which represent the significant research interests of investigators who are concerned with giving the field of social psychology theoretical structure and reporting research which is clearly focused, well designed, and competently conducted.

While social psychology is presently regarded by most as a field with indeterminate boundaries, it has as its central focus the investigation of the processes and products of social interaction at the interpersonal, intrapersonal, intergroup, and intragroup levels and the development of significant generalizations therefrom. In keeping with the more general meaning of the name of the journal emphasis will be placed on measurement of social behavior. However, this emphasis does not exclude the acceptability of good articles which must rely upon qualitative materials and analyses.

The editors and editorial consultants can be expected to subject manuscripts to rigorous criticism and screening according to the best standards of scientific research and at the same time avoid a sterile orthodoxy which would stultify the communication of creative ideas at the growing edge of the science. Thus the journal will strive to be flexible in its response to the publication needs of its contributors.

It is the intention of the editors to avoid any tendency toward professional provincialism and to invite contributions from any sector of the scientific community which promise to further the objectives of the journal.

Annual *subscription* rate $9.00; special rate to members of THE AMERICAN SOCIOLOGICAL SOCIETY $4.50. Single issues $2.50.

Specifications for Manuscript Preparation for Sociometry may be obtained from the Editorial Office.

Back Issues dated 1937 through 1955 should be ordered from the former publisher, Beacon House, Beacon, New York.

Notice of *change of subscriber's address* should be sent to THE SOCIETY four weeks in advance, indicating the old as well as the new address.

The Psychosocial Origins of Achievement Motivation [1]

BERNARD C. ROSEN, *University of Connecticut*
ROY D'ANDRADE, *Cambridge, Massachusetts*

The purpose of this study is to examine the origins of achievement motivation (*n* Achievement) within the context of the individual's membership in two important groups: family and social class. Specifically, this paper explores, through the observation of family interaction, the relationship between achievement motivation and certain child-training practices, and the relationship between these practices and the parent's social class membership.

The importance of group membership for personality development has been demonstrated many times. Perhaps the most important of these groups is the family, whose strategic role in the socialization process has led investigators to study the nexus between child-rearing practices and motivation formation. Thus, Winterbottom (15) examined the relationship between independence-mastery training and achievement motivation and found that achievement motivation is strongest among boys whose mothers (all of whom were middle class) expected relatively early indications of self-reliance and mastery from them.

Since many socialization practices are known to be dissimilar between social groups (3, 4), it might be expected that independence training practices would also differ. A study by McClelland *et al.* (8), later replicated by Rosen (10), demonstrated this to be the case: middle-class parents place greater stress upon independence training than lower class parents. The deduction from this finding that classes differ in their level of *n* Achievement was shown to be correct by Rosen (9) who found that, on the average, *n* Achievement scores for

[1] This paper is a report of an investigation supported by a research grant (M1495) to the senior writer from the National Institute of Mental Health, Public Health Service. The generous support, advice, and encouragement of David C. McClelland is gratefully acknowledged. The contributions of James Sakoda for his help on statistical procedures, Shirley Rosen for her work in administering the Thematic Apperception Test and in the preparation of the manuscript, and Marian Winterbottom for her work in scoring the TAT protocols are deeply appreciated. Roland Bonato, Diane D'Andrade, and June Schmelzer assisted in the collection of the data. David Bakun and Thomas Shipley made many helpful suggestions.

A special expression of thanks is due Superintendents Richard C. Briggs, Mansfield school system; Arthur H. Illing, Manchester school system; and Charles Northrup, Windham school system; Principals Stephen J. Ardel, Buchanan School; Harriet Atwood, Highland Park School; David J. DiSessa, Annie Vinton School; Esther Granstrom, Bowers School; Robert L. Perry, Natchaug School; Catherine Shea, Verplanck School; Lawrence Smith, Windham Center School.

middle-class adolescents were significantly higher than those for their lower class counterparts.

Significantly, although these studies flow logically from one another, in none of them were all three variables—group membership, child training practices, and *n* Achievement—studied simultaneously. Furthermore, there were certain gaps in these studies which called for theoretical and methodological modifications and additions. The nature of these gaps, and the contributions which it was the research objective of this study to make, are as follows:

Theoretical. The keystone around which studies of the origins of achievement motivation have been built is the notion that training in independent mastery is an antecedent condition of *n* Achievement (6, 15). This approach grew out of McClelland's and his associates' theory of the nature and origins of motivation. They argue that all motives are learned, that "they develop out of repeated affective experiences connected with certain types of situations and types of behavior. In the case of achievement motivation, the situation should involve 'standards of excellence,' presumably imposed on the child by the culture, or more particularly by the parents as representatives of the culture, and the behavior should involve either 'competition' with those standards of excellence or attempts to meet them which, if successful, produce positive affect or, if unsuccessful, negative affect. It follows that those cultures or families which stress competition with standards of excellence or which insist *that the child be able to perform certain tasks well by himself* . . . should produce children with high achievement motivation" (7).

Two distinctly different kinds of child-training practices are implicit in this theory. The first is the idea that the child is trained to do things "well"; the second, the notion that he is trained to perform tasks "by himself." The former has been called *achievement training* (2) in that it stresses competition in situations involving standards of excellence; the latter has been called *independence training* in that it involves putting the child on his own. The failure to disentangle these two concepts has resulted in a focus of attention upon independence training largely to the exclusion of achievement training, although the former is primarily concerned with developing self-reliance, often in areas involving self-caretaking (e.g., cleaning, dressing, amusing, or defending oneself). Although both kinds of training practices frequently occur together, they are different in content and consequences and needed to be examined separately. We believe that of the two training practices, achievement training is the more effective in generating *n* Achievement.

There is another component of independence training—one which is explicit in the idea of independence—that needed further exploration: *autonomy*. By autonomy, we mean training and permitting the child to exercise a certain amount of freedom of action in decision making. Although a related aspect of

autonomy—*power*—was studied by Strodtbeck (13), who examined the relationship between power distribution in the family, *n* Achievement, and academic achievement among a group of Jewish and Italian adolescents, no study had examined simultaneously the self-reliance and autonomy components of independence training. The operation of both components, we believed, tends to increase the power of independence training to generate *n* Achievement, since in itself high parental expectations for self-reliance may cause rebellion, feelings of rejection, or of apathy on the part of the child, while autonomy without parental expectations for self-reliance and achievement may be perceived as mere permissiveness or indifference.

In association with parental demands that the child be self-reliant, autonomous, and show evidence of high achievement, there must be sanctions to see that these demands are fulfilled. Winterbottom found that mothers of children with high *n* Achievement gave somewhat more intense rewards than mothers of children with low *n* Achievement. Little was known about the role of negative sanctions, or of the relative impact of sanctions from either parent. Further study was required of the degree and kind of sanctions employed by both parents to see that their demands are met.

Methodological. This study departed from two practices common in studies of the origins of *n* Achievement. The first practice is to derive data exclusively from ethnographic materials; the second, to obtain information through questionnaire-type interviews with mothers. Interviews and ethnographies can be valuable sources of information, but they are often contaminated by interviewer and respondent biases, particularly those of perceptual distortion, inadequate recall, and deliberate inaccuracies. There was a need for data derived from systematic observation of parent-child relations. It is not enough to know what parents *say* their child-rearing practices are; these statements should be checked against more objective data, preferably acquired under controlled experimental conditions, that would permit us to *see* what they do. In this study, experiments were employed which enabled a team of investigators to observe parent-child interaction in problem-solving situations that were standardized for all groups and required no special competence associated with age or sex.

An equally strong objection can be raised against the tendency to ignore the father's role in the development of the child's need to achieve. Apart from an earlier study of father-son power relations, no efforts had been made to determine the father's contribution to achievement and independence training—a surprising omission even granted the mother's importance in socializing the child in American society. Although we were not prepared to take a position on the nature of the role relationships between father, mother, and son with respect to this motive, we deliberately created experimental conditions which

would enable us to observe the way in which the three members of the family interacted in a problem-solving situation. Finally, this study incorporated in one design the variables of group membership, child-training practices, and motivation, variables that heretofore had not been studied simultaneously. In so doing we hoped to establish the nexus among class membership, socialization practices, and achievement motivation.

HYPOTHESES

This study was designed to provide data that would permit testing two basic hypotheses.

1. Achievement motivation is a result of the following socialization practices: (a) *achievement training*, in which the parents set high goals for their son to attain, indicate that they have a high evaluation of his competence to do a task well, and impose standards of excellence upon tasks against which he is to compete, even in situations where such standards are not explicit; (b) *independence training*, in which the parents indicate to the child that they expect him to be *self-reliant*, while at the same time permit him relative *autonomy* in situations involving decision making where he is given both freedom of action and responsibility for success or failure; (c) *sanctions*, rewards and punishments employed by parents to ensure that their expectations are met and proper behavior is reinforced. Although each contributes to the development of achievement motivation, achievement training is more important than independence training. Neither are effective without supporting sanctions.

2. Differences in the mean level of achievement motivation between social classes is in part a function of the differential class emphases upon independence and achievement training: middle-class parents are more likely than lower class parents to stress self-relience, autonomy, and achievement in problem-solving situations, particularly those involving standards of excellence. They are more likely to recognize and reward evidences of achievement, as well as to be more sensitive of and punitive toward indications of failure.

EXPERIMENTAL PROCEDURE

The subjects selected to provide data needed for the testing of these hypotheses about the origins of achievement motivation were 120 persons who made up 40 family groups composed of a father, mother, and their son, aged nine, ten, or eleven. The selection of the family groups began with testing the boy. Seven schools in three northeastern Connecticut towns were visited by the same field worker who administered a Thematic Apperception Test individually and privately to 140 boys, aged nine, ten, or eleven. As is customary in the TAT procedure, the subject was presented with a set of four ambiguous

pictures and asked to tell a story about each. His imaginative responses were then scored according to a method developed by McClelland and his associates which involves identifying and counting the frequency with which imagery about evaluated performance in competition with a standard of excellence appears in the thoughts of a person when he tells a brief story under time pressure. Experience has shown that this imagery can be identified objectively and reliably. It is the assumption of this test that the more the individual shows indications of evaluated performance connected with affect in his fantasy, the greater the degree to which achievement motivation is part of his personality (7). The stories were scored by two judges; the Pearsonian coefficient of correlation between scorers was .87, a level of reliability similar to those reported in earlier studies with this measure.

Subjects with scores of plus 2 to minus 4 (approximately the bottom quartile) were labeled as having low *n* Achievement, those with scores of plus 9 to plus 22 (approximately the top quartile) as having high *n* Achievement. Any boy with an I.Q. score below 98, with physical defects, whose parents were separated, or who had been raised during part of his life by persons or relatives other than his parents (e.g., grandparents) was eliminated from the sample.

Forty boys, matched by age, race, I.Q., and social class were chosen for further study. All were white, native born, and between nine and eleven years of age; the average was ten years. Half of the boys had high *n* Achievement scores, half had low scores. In each achievement motivation category, half of the boys were middle class, half were lower class. Their social class position was determined according to a modified version of the Hollingshead Index of Social Position (5) which uses the occupation and education of the chief wage-earner—usually the father—as the principal criteria of status. The middle-class father (class II or III) held either a professional, managerial, white-collar position or was self-employed as an owner of a small- to medium-size business. Often one or both parents in middle-class families were college graduates; all were high-school graduates. The parents of lower class (IV or V) boys were quasi-skilled or skilled workers in local factories, or owners of very small farms—often the farmers held factory jobs as well. Relatively few of these parents had completed high school, none had gone beyond high school.

It can be seen that the study was designed in such a way that the subjects fell into one of four cells, with the achievement motivation level of the boys and the class position of the parents as the classificatory variables. Within each cell there were ten families. This four-cell factorial design was constructed so as to facilitate the use of the analysis of variance technique in the statistical analysis of the data.

After the boy was selected, a letter was sent to his parents from the principal of the school asking their cooperation with the investigators. Later, appointments were made over the telephone to visit the families. Cooperation was very good; there were only two refusals. A pair of observers visited each family group, usually at night. There were two teams of observers, each composed of a man and woman. Both teams had been trained together to ensure adequate intra- and interteam reliability.

Once in the home, the observers explained that they were interested in studying the factors related to success in school and eventually to a career, and that the son was one of many boys selected from a cross-section of the community. When rapport had been established, the parents and their son were placed at a table—usually in the kitchen—and it was explained that the boy was going to perform certain tasks.

Experimental Tasks

The observers wanted to create an experimental situation from which could be derived objective measures of the parents' response to their son as he engaged in achievement behavior. Tasks were devised which the boy could do and which would involve the parents in their son's task performance. The tasks were constructed so that the subjects were often faced with a choice of giving or refusing help. At times they were permitted to structure the situation according to their own norms; at other times the experimenters set the norms. In some situations they were faced with decision conflicts over various alternatives in the problem-solving process. The observation of the parents' behavior as their son engaged in these experimental tasks provided information about the demands the parents made upon him, the sanctions employed to enforce these demands, and the amount of independence the child had developed in relations with his parents. A category system, similar to the Bales system (1), was devised to permit scoring interaction between parents and son so that the amount and form of each subject's participation could be examined. The investigators were able to learn from these interaction data how self-reliant the parents expected their son to be, how much autonomy they permitted him in decision-making situations, and what kind and amount of affect was generated in a problem-solving situation.

In creating the experimental tasks an effort was made to simulate two conditions normally present when boys are solving problems in the presence of their parents: (1) tasks were constructed to make the boys relatively dependent upon their parents for aid, and (2) the situation was arranged so that the parents either knew the solution to the problem or were in a position to do the task better than their son. In addition, tasks were created which tapped manual skills as well as intellectual capacities, although intelligence is a factor

in any problem-solving situation. It was for this reason that the experimenters controlled for I.Q.

In one particular respect the experimental situation was deliberately made atypical. The investigators sought to get the parents involved in the experiment by deliberately building stress into the situation. It was hoped that these *stress experiments* would so involve the parents that they would abandon their protective "company behavior" and generate more authentic action in several hours than could be gained through casual observation over several days. This manoeuvre was generally successful, although it is impossible to evaluate how and in what way the nature of the experiments and the presence of observers affected the subjects. It is a basic assumption of this study that by studying present-time interaction in a controlled situation one can achieve a valid picture of the patterns of interaction between parents and child most likely to have occurred in the child's earlier years. It is recognized, however, that the conflicting evidence about changes in socialization practices as the child grows older leaves the wisdom of this assumption an open question.

Pretesting had shown that no single task would provide sufficient data to test all hypotheses. Hence, five tasks were constructed, each designed to attack the problem from a somewhat different angle and yet provide certain classes of data that could be scored across tasks. The five tasks used in this study are as follows:

1. *Block Stacking*. The boys were asked to build towers out of very irregularly shaped blocks. They were blindfolded and told to use only one hand in order to create a situation in which the boy was relatively dependent upon his parents for help. His parents were told that this was a test of their son's ability to build things, and that they could *say* anything to their son but could not touch the blocks. A performance norm was set for the experiment by telling the parents that the average boy could build a tower of eight blocks; they were asked to write down privately their estimate of how high they thought their son could build his tower. The purposes of this experiment were (a) to see how high were the parents' aspirations for and evaluations of their son, e.g., if they set their estimates at, above, or below the norm; (b) to see how self-reliant they expected or permitted their son to be, e.g., how much help they would give him.

There were three trials for this task. The first provided measures of parental evaluations and aspirations not affected by the boy's performance; the second and third trial estimates provided measures affected by the boy's performance. The procedure for the third trial differed from the first two in that the boy was told that he would be given a nickel for each block he stacked. Each member of the family was asked to estimate privately how high the boy should build his tower. No money would be given for blocks stacked higher than the

estimate nor would the subject receive anything if the stack tumbled before he reached the estimate. Conservative estimates, hence, provided security but little opportunity for gain; high estimates involved more opportunity for gain but greater risk. The private estimates were then revealed to all and the family was asked to reach a group decision. In addition to securing objective measures of parental aspiration-evaluation levels, the observers scored the interaction between subjects, thus obtaining data as to the kind and amount of instructions the parents gave their son, the amount of help the son asked for or rejected, and the amount and kind of affect generated during the experiment.

2. *Anagrams.* In this task the boys were asked to make words of three letters or more out of six prescribed letters: G, H, K, N, O, R. The letters, which could be reused after each word was made, were printed on wooden blocks so that they could be manipulated. The parents were given three additional lettered blocks, T, U, and B, and a list of words that could be built with each new letter. They were informed that they could give the boy a new letter (in the sequence T, U, B) whenever they wished and could say anything to him, short of telling him what word to build. There was a ten-minute time limit for this experiment. Since this is a familiar game, no efforts were made to explain the functions of the task.

The purposes of this experiment were: (a) to see how self-reliant the parents expected their son to be, e.g., how soon they would give him a new letter, how much and what kind of direction they would give him, if they would keep him working until he got all or most of the words on the list or "take him off the hook" when he got stuck. And (b) to obtain, by scoring interaction between the subjects, measures of the affect generated by the problem-solving process, e.g., the amount of tension shown by the subjects, the positive and negative remarks directed toward one another.

3. *Patterns.* In this experiment the parents were shown eight patterns, graduated in difficulty, that could be made with Kohs blocks. The subjects were informed that pattern 1 was easier to make than pattern 2, pattern 3 was more difficult than 2 but easier than 4, and so forth. The subjects were told that this was a test of the boy's ability to remember and reproduce patterns quickly and accurately. Each parent and boy was asked to select privately three patterns which the boy would be asked to make from memory after having seen the pattern for five seconds. All three patterns were chosen *before* the boy began the problem solving so that his performance in this task would not affect the choice of the patterns. Where there were differences of choice, as inevitably there were, the subjects were asked to discuss their differences and make a group decision. Insofar as possible the observers took a verbatim account of the decision-making process, scoring for three kinds of variables: (a) the number of acts each subject contributed to the decision-making process,

(b) the number of times each individual initiated a decision, and (c) the number of times each subject was successful in having the group accept his decision or in seeing to it that a decision was made.

The purposes of this experiment were: (a) to obtain another measure of the parents' evaluations of and aspirations for the boy, e.g., whether they would pick easy or difficult tasks for him to do; (b) to get a measure of the autonomy permitted the boy, e.g., whether they would let him choose his own patterns or impose their choices upon him; and (c) to see how much help they would give him and what affect would be generated by the experiment.

4. *Ring Toss.* In this experiment each member of the group was asked to choose privately ten positions, from each of which the boy was to throw three rings at a peg. The distance from the peg was delineated by a tape with 1-foot graduations laid on the floor. The subjects were told that this was a test of discrimination and judgment and that after each set of three tosses they would be asked to make a judgment as to the best distance from which to make the next set of tosses. Group decisions were made as to where the boy should stand. The purposes of this experiment were: (a) to see whether the parents imposed standards of excellence upon a task for which no explicit standard had been set, e.g., whether the parents would treat this as a childish game or see it as a task which could and should be done well. Would they choose easy or difficult positions? (b) To determine how much autonomy they permitted their son, e.g., would they let him choose his own position?

5. *Hatrack.* The Maier Hatrack Problem was used in this experiment. The boy was given two sticks and a C-clamp and instructed to build a rack strong enough to hold a coat and hat. His parents were told that this was a test of the boy's ability to build things. In this task no one was given the solution at the beginning of the experiment. For the first time the parents had no advantage over the boy—a most uncomfortable position for many parents, particularly the fathers. This stress situation was created deliberately to maximize the possibility of the problem generating affect, as was often the case, with some hostility being directed at the observers. After seven minutes the parents were given the solution to the problem. The purposes of this experiment were: (a) to see how self-reliant the parents expected their son to be. After receiving the solution what kind of clues would the parents give the boy? How hard would they expect him to work on his own? (b) To obtain measures of the affect created in an unusually frustrating situation. How would the parents handle their frustration? Would they turn it against the boy?

Category System

References have been made to the use of a category system for scoring interaction between subjects. A brief description of this system, shown in Diagram

1, is in order. Most of the subjects' verbal and some of their motor behavior (e.g., laughing, hand-clapping, scowling) was scored in one of twelve categories. In eight of these categories were placed acts involving relatively strong affect. Four additional categories were used to distinguish between various kinds of statements—either giving, requesting, or rejecting directions—which contained very little or no affect. A distinction was made between negative and

DIAGRAM 1

The System of Categories Used in Scoring Parent-Child Interaction

+X	Expresses approval, gives love, comfort, affection
+T	Shows positive tension release, jokes, laughs
+E	Gives explicit positive evaluation of performance, indicates job well done
+P	Attempts to push up performance through expression of enthusiasm, urges cheers on
N	Gives nonspecific directions, gives hints, clues, general suggestions
S	Gives specific directions, gives detailed information about how to do a task
aa	Asks aid, information, or advice
ra	Rejects aid, information, or advice
−P	Attempts to push up performance through expressions of displeasure; urges on, indicating disappointment at speed and level of performance
−E	Gives explicit negative evaluation of performance, indicates job poorly done
−T	Shows negative tension release, shows irritation, coughs
−X	Expresses hostility, denigrates, makes sarcastic remarks

positive affective acts. Affective acts associated with explicit or implicit evaluations of the boy's performance which aimed at motivating or changing his behavior were scored differently from affective acts which involved reactions to the boy and only indirectly to his performance.

Directional acts by the parents were remarks designed to help the boy perform his task. A distinction was made between *specific* directions (S) which were acts instructing the subject to do particular things which would facilitate task completion, and *nonspecific* (N) which were acts aimed at giving the subject some information but not specific enough to enable him to rely entirely upon it. It was believed that nonspecific statements were more likely than specific statements to create self-reliance in the child.

The affective acts were schematized in two sets—one positive, the other negative. The first set was comprised of acts involving direct expressions of emotions toward another person, not necessarily in the context of task performance, either of a positive character (+X), such as expressions of love or approval, or of a negative character (—X), such as indications of hostility and rejection. Another set was of acts involving release of tension, either associated with positive affect (+T) such as grins, laughter, jokes, or negative affect (—T) such as scowls, coughs, or irritated gestures. Tension-release acts differ from acts of direct emotion (X) in that the former were not focused toward any person but were diffused, undirected reactions to the general situation. The next set of acts involved parental evaluation of the boy's performance. Those acts in which the parents stated that the boy was doing the task well were scored as positive evaluations (+E), while statements that the boy was doing poorly were scored as negative evaluations (—E). The last two categories involved acts aimed at urging or pushing the boy to perform more effectively. These "pushing up the performance level acts" were scored in one of two categories. Those acts in which the parents "cheered" the boy on while at the same time indicating that they expected him to do better were scored as positive pushing acts (+P); negative pushing acts (—P) were statements in which the parents sought to improve the boy's performance by indicating in a threatening way that they thought he could do better.

Only four kinds of acts were scored for the boy: whether he asked for aid (aa), rejected aid (ra), showed positive tension (+T) or negative tension (—T). An act was defined as the smallest segment of verbal or motor behavior which could be recognized as belonging to one of the twelve categories in the system. The actor rather than the target of the acts was used as the observer's frame of reference.

This category system involves a good deal of inference on the part of the observer—a factor which can make for low observer reliability. To ensure adequate reliability, the interaction in each family was scored by two observers who had been trained to work as a team. In the early stages of the field work tape recordings were taken of family interaction. After the experiments the observers rescored the interaction protocols and discussed scoring differences in order to increase interobserver reliability. Tape recordings were discontinued when the scorers felt that their scores were substantially the same. Each team of observers visited 20 families. The reliability of observers for the gross number of acts scored is high. The Pearsonian coefficient of correlation between the first pair of scorers is plus .93, and for the second pair plus .97. No significant differences between pairs of observers has been discovered.

EXPERIMENTAL FINDINGS

The fundamental hypothesis of this study is that achievement motivation develops out of repeated affective associations with two kinds of experiences: achievement training and independence training. It was hypothesized that high *n* Achievement scores were most likely to be found among children whose parents (1) gave them *achievement training* (i.e., imposed standards of excellence upon tasks, had high expectations for their son's achievement, and indicated that they evaluated highly his competence to do a task well), (2) who gave them *independence training* (i.e., expected high evidences of self-reliance while at the same time permitting relative autonomy in decision making), and (3) employed sanctions to reinforce appropriate behavior. Data to test these hypotheses were derived from an examination of parental estimates and choices and from the observation of parent-child interaction. From these data several measures of achievement and independence training and sanctions were constructed. We turn now to an examination of the manner in which these measures were derived and their relationship to achievement motivation.

Achievement Training and Achievement Motivation

Measures of achievement training were obtained from the estimates and choices made by the parents in three tasks: Block Stacking, Patterns, and Ring Toss. Each task provided measures of achievement training which, though positively related to one another, were sufficiently independent to require their being treated as separate scores. This situation was a result of intended fundamental differences between the tasks. Thus, in one task (Block Stacking) the parents were asked how well their son would do, in a second situation (Patterns) they were asked to choose a task for him to perform, while on the last problem (Ring Toss) they were given on open-ended situation which they could structure in a number of ways.

Parental Aspirations and Evaluations

The parents' estimates of how well their son would do in the Block Stacking task are considered measures of their aspiration for and evaluations of him. In this case the estimates were made against a stated norm. The parents' first estimate, unaffected by any previous performance in this task, is conceived to be primarily a measure of parental aspirations for the boy. Undoubtedly, this estimate is somewhat affected by the parents' evaluation of the boy's competence, for aspiration and evaluation levels are often intricately mixed. Nonetheless, since presumably the parents had never seen their son perform a block-stacking test with one hand while blindfolded, we believe the element of aspiration level to be dominant in this measure.

We began by comparing the fathers of boys with high *n* Achievement with

the fathers of boys with low *n* Achievement, and the mothers of high *n* Achievement boys with the mothers of boys with low *n* Achievement. It had been predicted that the parents of the boys with high *n* Achievement scores would give higher estimates than parents of boys with low scores. And in fact the data are in the predicted direction; the fathers and mothers of boys with high *n* Achievement scores on the average give higher estimates, but the differences are not statistically significant. However, when father's and mother's scores are summed together the differences between parental groups is significant ($F=4.09$, $P<.05$).

Parental estimates for the second and third trials of this task were combined to form a single score. Since the means and standard deviations of both estimates were about the same, no transformation into standard scores was necessary. Both estimates had been influenced considerably by the boy's performance. The combined score was considered as a measure of the parents' aspiration-evaluation of the boy as affected by his performance against a given standard of excellence. The data show that the mothers of boys with high *n* Achievement scores give considerably higher estimates for the second and third trials of this task than do the mothers of boys with low scores ($F=10.28$, $P<.005$). The differences between the fathers, although in the predicted direction in that the fathers of boys with high *n* Achievement scores tend to give higher estimates, are not statistically significant.

The Patterns task was designed to provide additional and supplemental measures of parental aspiration-evaluation levels. In this situation the parents were not asked to estimate how well their son would do, but actually to select three tasks, graded in difficulty, for their son to perform. Since all three choices were made before their son began the problem, these scores could be considered as performance-free measures of the parents' evaluation of the boy's capacity to do the task, plus their aspiration for him to do the more difficult task in a situation structured by degree of difficulty with no stated norm or group average.

Each subject's three choices were combined by simple addition (no transformation of scores was necessary) to provide a single score for each person. We had expected that the parents of boys with high achievement motivation scores—in keeping with their tendency to have higher aspirations for and evaluations of their boys—would choose the more difficult patterns for their sons. The data indicate a difference in this direction, but the differences are small and not significant for either fathers or mothers.

Parental Standards of Excellence

In the two experiments just described some parents would very likely have imposed standards of excellence upon the tasks even if the experimenters had

not done so. However, since they had been asked to make their estimates or choices in situations where standards were explicit (as in the Block Stacking task where a group performance norm had been given, or in the Patterns task where the complexity of the patterns had been clearly graded by the experimenters) it could not be clearly seen from these experiments whether parents differed in their tendencies to impose standards upon the problems their children were expected to solve. The Ring Toss experiment was devised for this purpose. In this experiment no norm or group standard of excellence was set by the investigators. Each parent was asked to make ten choices of "the best place for your son to stand." After each choice the boy threw three rings at the peg. A measure of the height of the standard of excellence each parent sets for the boy was derived by summing the choices (number of feet the boy is asked to stand from the peg) of each parent.

We hypothesized that parents who imposed standards of excellence upon this normless task would be most likely to structure tasks generally in these terms. Our expectation was that the parents of boys with high motivation scores would be more likely to impose standards of excellence upon this task in that they would place their sons farther from the peg than would parents of boys with low scores. The data tend to support this expectation: the fathers and mothers of high *n* Achievement boys, on the average, chose positions further from the peg than the parents of children with low achievement motivation. The differences, however, are significant only in the case of the mothers ($F=5.47$, $P<.025$). Combining the scores for fathers and mothers increases the differences between parental groups ($F=6.99$, $P<.01$).

The Anagrams experiment involved another task for which the investigators had set no explicit standard of excellence, the parents and boy merely being provided with lettered blocks out of which words could be made. This experiment had been designed primarily to provide measures of independence training. The sum of the time at which new letters (the parents shared three letters) were given by both (or either) parents was treated originally as a measure of self-reliance training, i.e., the longer the parents delayed in giving the boy new letters the more indication that they expected him to work longer and harder at a problem on his own. This experiment revealed a clear difference between parental groups but not in the direction we had predicted: the parents of high *n* Achievement boys gave new letters *sooner* than the parents of boys with low *n* Achievement ($F=6.28$, $P<.025$).

This finding, so different from what had been expected, prompted a re-evaluation of the task and a further (albeit ex post factum) interpretation of the data. We believe, now, that we were mistaken in assuming that this task would only measure self-reliance training. Rather, our observations indicate that this experiment elicited from some parents a type of achievement training

in which the element of competition with a standard of excellence was very strong. The parents (especially the mothers) of boys with strong achievement motivation, it appeared, tended to perceive the task not so much as one which their son should do on his own, *but as a challenge to do well.* They reacted to this experiment with more emotion and competitiveness than displayed toward any other problem. The mother, in particular, was eager for her son to do well and often became anxious when he stopped making words. Both parents typically showed keener pleasure and disappointment at the boy's success or failure than was ordinarily displayed by the parents of boys with weak achievement motivation. The boy with strong motivation tended to receive letters sooner, we believe, because his parents were eager to see him make words and because of their reluctance to frustrate him to a point where his motive to excel in this important area would be destroyed.

In choosing to structure this task in terms of achievement training rather than self-reliance training—a decision congruent with our theoretical position that the former is the more important in the development of achievement motivation, the parents of high *n* Achievement boys are nonetheless exhibiting less self-reliance training than the parents of low *n* Achievement boys. This is contrary to our theoretical expectations, but is similar, as we shall see below, to other empirical data which indicate that the mothers of high *n* Achievements score low on self-reliance training—and typically the mother played her most prominent role in the Anagrams experiment.

Achievement Training and Performance Levels

The behavior of people with high achievement motivation is characterized by persistent striving and general competitiveness. It would follow from this, other things being equal, that boys with high achievement motivation would perform better than those with low motivation—and in fact this proved to be the case. Boys with high *n* Achievement tend to build higher towers of blocks, construct patterns faster, and make more words in the Anagrams task. The differences are significant in the case of the Block Stacking task ($F=8.16$, $P<.005$), but not for the Patterns and Anagrams experiments. In the latter two tasks the individual's performance is very greatly affected by his intelligence. Since I.Q. score was one of the variables controlled in this study (there were no differences between the mean I.Q. scores of boys with high *n* Achievement and their peers with low achievement motivation), even these small differences are surprising. The superior performance of high *n* Achievement boys appears to be more a function of greater self-reliance and zest in competitive activity than of intelligence. Thus, boys with high achievement motivation tend to ask for less aid (aa), are more likely to reject offers of help from their parents (ra), and appear to get more pleasure out of participating in the

experiments—they show less evidence of negative affect (—T) and more of positive feelings (+T). Although for only one of these variables—asks aid ($F=5.76$, $P<.05$)—is there a significant difference between groups, the direction of the differences in all four cases consistently points to greater self-reliance and self-assurance on the part of boys with high need for achievement.

The question arises, then, whether the higher estimates of the parents of boys with high *n* Achievement are not a natural response to his superior performance, rather than a measure of their aspirations for and evaluations of him. In two tasks, Block Stacking and Anagrams, where there were significant differences between parental groups, the analysis of covariance technique was used to control the effect of the boy's performance upon parental estimates. This technique enabled us to compute what the parents' estimates of the boy's performance, or their willingness to give him a new letter, would be if the performance of high and low *n* Achievement boys were made comparable. An analysis of covariance of the mother's estimates for the second and third trials in the Block Stacking task—there had been no significant difference between fathers—provided an adjusted F ratio for the mother's estimates which, although reduced, is still statistically significant ($F=6.17$, $P<.05$).

In order to see whether the speed with which the parents give their sons new letters in the Anagrams task is merely a function of the rate at which the boy builds words—it will be remembered that high *n* Achievement boys made more words during the ten-minute time limit for this task than boys with low achievement motivation—an analysis of covariance was computed with the number of words controlled. The adjusted F ratio not only remained significant but increased very slightly from 6.28 to 6.45, both ratios being significant at the .05 level.

In the Ring Toss experiment there was no question of whether superior performance by high *n* Achievement boys had influenced their parents to make significantly higher choices, because in this task high *n* Achievement boys were less successful in placing rings around the peg than their low *n* Achievement peers. The reason for this is simple: the parents of high *n* Achievement boys tended to place their sons farther away from the peg and consequently the number of their successes was smaller; the tetrachoric correlation between the number of successes and the distance away from the peg is —.31.

These data clearly indicate that the parents were not responding *merely* to the boy's performance, but this is not to say his ability had no influence on their estimates and choices. In the Block Stacking experiment, for example, the parents' aspirations and evaluations tend to go up as the boy's performance improves. This is particularly true for the mothers of high *n* Achievement boys, who are more affected by the boy's performance than the fathers.

Thus, for the first trial of the Block Stacking task the F ratio for fathers of high *n* Achievement boys is larger than mothers, but for the second and third trials the mothers' F ratio jumps from 1.24 to 10.12, while the fathers' F increases only slightly, from 2.78 to 3.45. This is particularly interesting in view of the fact that fathers were more active in this task than mothers, who seemed to feel that this was the sort of task about which males were more informed. Nonetheless, even though the mothers of high *n* Achievement boys did not generally perceive this task as falling under their area of parental training, they responded quite strongly to their sons' performance in their aspirations and evaluations. From a longer developmental point of view, of course, these data do not permit us to state with certainty as to which came first, the higher aspiration-evaluation levels of the parents, or the boy's superior performance; the data strongly suggest that they very probably interact.

INDEPENDENCE TRAINING, SANCTIONS, AND ACHIEVEMENT MOTIVATION

Earlier we distinguished between achievement training and independence training; the latter was broken down into two components: self-reliance training and the granting of relative autonomy in decision making. Associated with *both* independence and achievement training are sanctions—rewards and punishments—administered by the parents to reinforce appropriate behavior in the child. The data to index these variables were obtained by examining the interaction between parents and child as they engaged in the experimental tasks, and by observing the decision-making process in those instances where the subjects were asked to make a group estimate or choice of what the boy should do.

The scoring of parent-child interaction produced a voluminous amount of data: hundreds of acts were scored during the average three-hour experimental session. It was necessary, therefore, to combine scores where possible. Fathers' scores for each category of action were combined across all tasks (not including specific directions), except for the Ring Toss task where only the decision-making process had been scored. Tetrachoric intercorrelations between all scores were computed and a first step cluster analysis using James Sakoda's modification (11) of Tryon's Profile method (14) was employed to obtain clusters. Those scores which clustered together were then transformed into standard scores and added together. Scores for two types of acts—negative evaluation of the boy (—E) and direct expression of positive feelings (+X)—were eliminated because of the very low frequency of acts in these categories.

Three clusters resulted from this analysis which appear to be conceptually and theoretically meaningful:

1. *Warmth.* A combination of positive tension (+T) and positive evaluational acts (+E), indicating generally pleasant, happy, anxiety-relieving,

laughing-joking behavior. This cluster provides a measure of the amount of positive affect the parents put out while the boy is working.

2. *Rejection.* A combination of negative tension (—T) and hostile acts (—X), indicating generally unpleasant, irritated, unhappy behavior. This cluster provides a measure of the negative affect the parent gives out while the boy is working.

3. *Pushing.* A combination of friendly (+P) and irritated (—P) statements by the parents urging the boy to work hard, to get on with the task. This cluster provides a measure of one kind of pressure parents put on the boy to meet their expectations.

The acts which occurred most frequently were directional statements: specific statements (S) and nonspecific statements (N). Intercorrelations between father's S scores for each task were computed and were sufficiently high between the Block Stacking, Anagrams, and Hatrack tasks to permit their combination. S scores for Patterns proved to be relatively unrelated to the other S scores and hence were not used. Scores were first transformed into standard scores before they were combined. Intercorrelations between father's nonspecific statements (N) for each task were computed. The intercorrelations were sufficiently high to permit combining all nonspecific statements across tasks. Transformation to standard scores was not necessary. Intercorrelations were then computed between all interaction scores (S, N, P, Warmth, Rejection). All intercorrelations are low, indicating that these measures are relatively independent and can reasonably be treated as distinct variables.

In order to see if the mother's scores could be combined in the same way as the father's scores, correlations between those scores which had been combined for father were computed for mother. No large discrepancies between father's intercorrelations and mother's intercorrelations were discovered (with one exception—the S scores), and the mother's scores were combined to form the same clusters that had been derived for the father. Mother's S scores do not intercorrelate in the same way as do father's, probably because there is a sex differential in parental participation in the tasks. Fathers tend to speak most during the Block Stacking, Ring Toss, and Hatrack experiments; mothers speak most frequently during the Anagrams and Patterns experiments. Mother's S scores for Anagrams, Hatrack, and Ring Toss were combined; S scores for Block Stacking, lowly correlated with other S scores, were not used. No score transformation was necessary. Intercorrelations between all mother's interaction scores were computed. Most intercorrelations are low but on the whole correlations are higher for mother than for father, hence these variables are less independent for mother than for father. This is probably because mothers tend to speak less than fathers. Generally, the greater the frequency of acts, the more the different types of acts are independent of each other.

Self-reliance Training

Directional statements (S and N) were used as measures of self-reliance training; a large number of N type acts in which the child is *not told exactly* what to do, but merely given a clue would indicate independence training; while a high incidence of specific directions (S) would be a sign of little such training. We had predicted that the parents of high need Achievement boys would generate more "N" acts and fewer "S" acts than the parents of boys with low achievement motivation. The data show that the fathers and mothers of high *n* Achievement boys give out *proportionately* more "N" acts and fewer "S" acts, but the differences are not significant. For total number of acts, the fathers of high *n* Achievement boys score higher on "N" acts and lower on "S" acts; the mothers of high *n* Achievement boys also score lower on "S" acts, but unlike the fathers they also score lower on "N" type acts. Although in all but one case—the lower N score for mothers of high *n* Achievement boys—the differences are in the direction predicted, in none of these cases are the differences statistically significant.

Pushing statements (+P and —P) are acts aimed at motivating the boy to work harder. An argument can be made for considering this type of act as reflecting either independence *or* achievement training, but in the context in which we observed this type of behavior it seemed to us to be primarily an index of independence training. We believe that a large number of pushing statements indicates *low* self-reliance training, for such statements often came from parents who appeared unable to sit back and let their boy work at his own speed but felt impelled to make him do well by urging or shouting him on. This type of parent seemed to assume that the boy had little internal need to excel, and that without external pressure he would soon run out of gas. As had been predicted, the fathers of high *n* Achievement boys gave fewer pushing statements than the fathers of boys with low achievement motivation. However, the reverse was true for the mothers: in this case it was the mothers of high *n* Achievement boys who had the higher score for pushing type statements. In none of these cases, however, are the differences statistically significant.

Autonomy

An index of the autonomy permitted the boy in decision making, another aspect of independence training, was derived by observing the family decision-making process in three tasks: Block Stacking (third trial), Patterns, and Ring Toss. The observers scored three types of behavior in this process. (1) The number of acts each subject contributed to the decision making, (2) the number of times each subject initiated the decision process by being the first to present a choice for consideration, (3) the number of times an individual made

the decision for the group or stated the final judgment. In all these decision-conflict situations somebody must state the final resolution, whether authoritatively or only as the summation of consensus. This final summation or statement was scored as "deciding." Intercorrelations between these three types of acts were computed; all correlations were sufficiently high to permit their being combined into a single score labeled "autonomy." The transformation of raw scores into standard scores was not necessary. It should be remembered that in this experiment the lower the parental autonomy score, the greater the autonomy permitted the boy.

We expected that the parents of high need Achievement boys would give their son more autonomy in the decision-making process than would be granted boys with low *n* Achievement. The data are in the direction predicted with respect to the fathers, but quite the reverse is true in the case of the mothers: the mothers of low *n* Achievement boys tend to grant greater autonomy to their sons. Unfortunately, none of these differences is statistically significant.

Sanctions

Typically, positive and negative reinforcements are associated with any learning situation—rewards for success and punishment for failure. We had predicted that the parents of boys with high achievement motivation would score higher an Warmth (positive affect) and lower on Rejection (negative affect) than the parents of low *n* Achievement boys. *The data show that the mothers of high* n *Achievement boys score significantly higher on Warmth than the mothers of low* n *Achievement boys* ($F=8.87$, $P<.01$). The differences between fathers, although in the predicted direction, are not significant ($F=4.13, P<.10$).

Fathers of boys with high *n* Achievement tend to score lower on Rejection. The reverse is true for the mothers: *the Rejection scores are higher for mothers of high Achievement motivation boys than for the mothers of low* n *Achievement boys*. None of these differences in Rejection, however, are statistically significant.

Parental Profiles and Motivation

In the analysis of data so far the relationship of each variable to achievement motivation was examined separately. The Split Plot type of analysis of variance was next employed to permit the examination of all variables simultaneously for each parent. All scores were transformed into standard scores, and in the case of four variables (S, Pushing, Rejection, and Autonomy) where low parental scores for parents had been hypothesized as producing high *n* Achievement the direction of the scores was reversed. Hence, we have labeled

these as "Fewer Specific Statements," "Less Pushing," "Less Rejection"; a high score for Autonomy means that the parents have given their son a relatively high amount of autonomy. This was done in order to make it possible to sum across all variables and arrive at a meaningful figure for each parent. A mean score for each variable for fathers and mothers was computed and the distance from the mean in standard deviations was plotted; the profiles for fathers are shown in Chart 1, for mothers in Chart 2.

CHART 1

Profiles for Father: High n *Achievement Group and Low* n *Achievement Group*

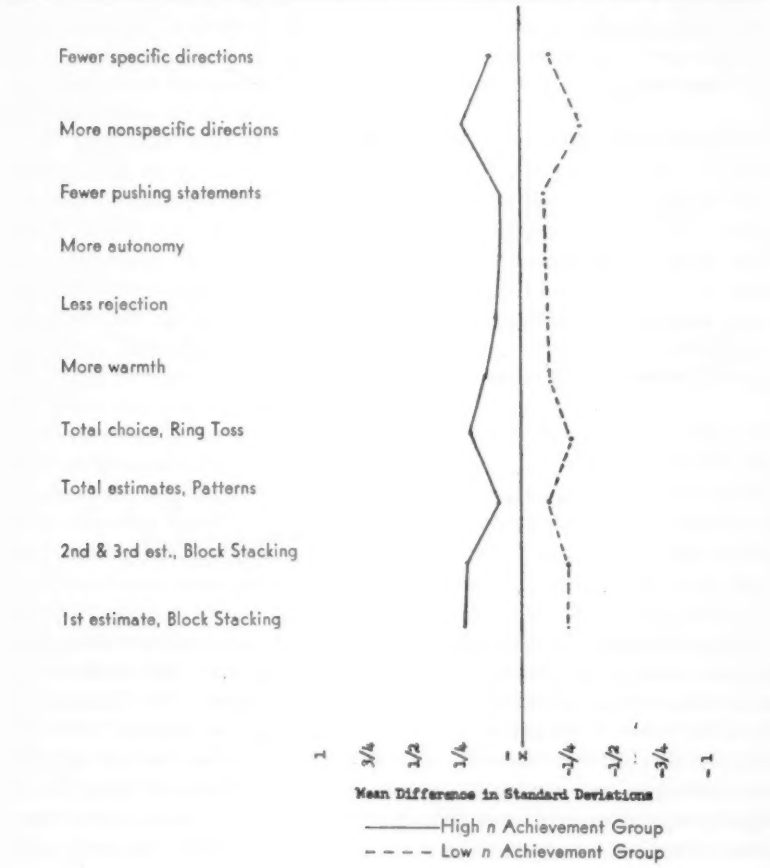
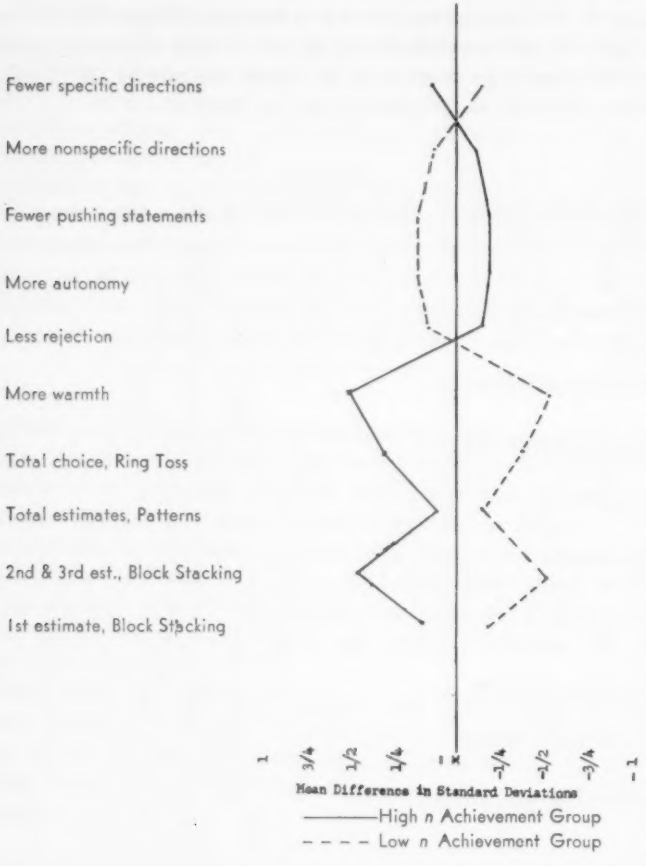


CHART 2

Profiles for Mothers: High n *Achievement Group and Low* n *Achievement Group*



A Split Plot analysis of variance reveals that there are significant differences *in levels* between the profiles of the parents of boys with high achievement motivation and the parents of boys with low *n* Achievement. The difference in levels for fathers is greater ($F=10.09$, $P<.005$) than for mothers ($F=4.77$, $P<.05$). By difference in level we mean that when the scores for each variable are summed for each parent, the parents of high *n* Achievement boys have a significantly higher *total* score than the parents of low *n* Achievement boys. Thus, although some variables are not significant when tested separately, and

others only barely significant, when the scores of all the variables are pooled together each contributes something to the total variance and the result is a significant difference between groups. This is most apparent in Chart 1 where the fathers' profiles are compared. It can be seen that the mean scores for the fathers of boys with high achievement motivation are higher for every variable, where "highness" was predicted as being positively related to high *n* Achievement. It should be remembered that the scores for S, P, Rejection, and Autonomy were reversed so that what appears as a "high" score in the chart is in fact a low score. In the case of these four variables, low parental scores were predicted as tending to produce high *n* Achievement. The difference between the two groups of fathers is not great, it is the *consistency* of the direction of these differences which when summed together make for a significant difference between the two groups. Another way of saying this is to point out that since the profiles of the two groups are almost parallel, there is no significant interaction between *n* Achievement groups and test variables when the level differences are taken out of the profiles.

The profiles for mothers are not parallel so that even when level differences are taken out the profiles for the mothers of high *n* Achievement boys remains significantly different from that of the mothers of low *n* Achievement boys ($F=2.30$, $P<.025$). The fact that there are interaction differences for mothers is not surprising when we remember that some of the variables which discriminate between groups for fathers do not do so for mothers, while other variables which are not significant for fathers are highly significant for mothers (e.g., Warmth). There are even some reversals in that mothers of high *n* Achievement boys give fewer nonspecific directions, more pushing statements, and are more dominant than the mothers of low *n* Achievement boys.

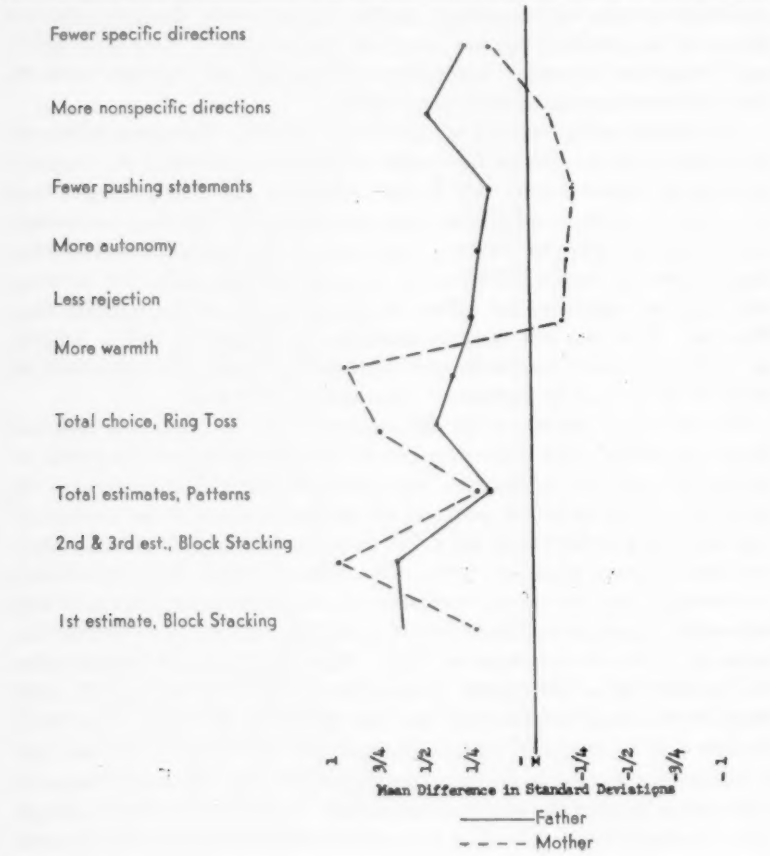
Another way of examining the differences between the parents of high and low *n* Achievement boys is shown in Chart 3. On this chart there is a profile for fathers and a profile for mothers. Each profile is obtained by computing the mean difference between the scores of the fathers, and then of the mothers, of high and low *n* Achievement boys. For example, for the first estimate Block Stacking the average standard score of the fathers of low *n* Achievement boys is subtracted from the average standard score of the fathers of boys with high achievement motivation. The difference is positive, that is, on the average the father of a high *n* Achievement boy had a higher score than the average father of a boy with low *n* Achievement. If the difference had been negative the point would have been plotted below the zero line, indicating that the average father of a low *n* Achievement boy had made the higher estimate. If there had been no difference, the point would have been on the zero line; the farther the point is plotted away from the zero line, the greater the difference between groups. Thus, for mothers with respect to the variable Warmth the difference between

n Achievement groups is more than plus ¾ indicating that the average score for the mothers of high *n* Achievement boys is more than three quarters of a standard deviation higher than that of the mothers of low *n* Achievement boys.

Chart 3 shows that the fathers of high *n* Achievement boys when compared with the fathers of low *n* Achievement boys tend to give higher estimates and choices for the Block Stacking and Patterns tasks, place their sons farther away from the peg in the Ring Toss experiment, are more warm, less reject-

CHART 3

Profiles for Fathers and Mothers: Mean Difference between Fathers of "Highs" and Fathers of "Lows"; Mothers of "Highs" and Mothers of "Lows"



ing, give their boy more autonomy, are less pushing, and give more nonspecific directions and fewer specific ones. The two groups of mothers when compared with one another present a somewhat different picture. The mothers of high *n* Achievement boys also tend to give higher estimates and choices for the Block Stacking and Patterns tasks, place their sons farther away from the peg in the Ring Toss experiment, give fewer specific directions, and are more warm. However, they are more rejecting, give the boy less autonomy, are more pushing, and give fewer nonspecific directions.

Parental Role Relationships and Achievement Motivation

The relationships of achievement motivation to certain training practices employed by *fathers and mothers considered separately* have now been described. But we know that the boy is subject to the influence of both parents, and that he belongs to a family unit in which the expectations and role behavior of both parents may have important consequences for his motivation. We know also that the parental pairs in this study did not always have the same expectations of the boy or the same reactions to his performance. The question then arises: Is there a difference between the parents of high and low *n* Achievement boys *when the parents are considered as a family pair?* Specifically, are the scores of the parents of high *n* Achievement boys more likely to be similar or dissimilar to one another than are the scores of the parents of boys with low achievement motivation? And what are the consequences of these similarities or dissimilarities for the development of achievement motivation?

The correlation technique made it possible to determine the relationship between the scores of fathers and mothers *considered as pairs*. Two sets of Pearsonian coefficients of correlation were computed for each variable: one set for the fathers and mothers of boys with high motivation, the other for the parents of the low motivation group. These correlations are taken as measures of the role relationships between parental pairs. A positive correlation is considered as indicating a *similar* role relationship; i.e., the parental scores tend to be high or low together. A negative correlation is considered as indicating a *dissimilar* role relationship, with parental scores moving in opposite directions. It must be stressed that these correlations tell us nothing about how high or low the scores are, merely that they are similar or dissimilar. To discover if there were any significant differences between the parents of high and low *n* Achievement boys the correlation coefficients were transformed into Z scores and the Standard Error of Difference computed. The ratio of the difference between Z's and the Standard Error of Difference was then computed and tested for significance. Table 1 presents the correlation for each measure, first for parents of high *n* Achievement boys and second for parents

of low *n* Achievement boys. The probability of difference between these two groups is shown in column 3.

Achievement motivation has been defined in terms of affective arousal over evaluated performance in connection with standards of excellence. Such standards are imparted to the individual typically by the parents, who indicate that he is expected to perform well in relation to these standards of excellence.

TABLE 1

Correlation Analysis of Pairs of Parents of High n *Achievement Boys Compared with Pairs of Parents of Low* n *Achievement Boys*

	F x M (Highs)	F x M (Lows)	P
Achievement Training			
1st estimate, Block Stacking	.50	.36	.308
2nd estimate, Block Stacking	.57	.31	.171
3rd estimate, Block Stacking	.36	.35	.480
1st choice, Patterns	.36	−.09	.087
2nd choice, Patterns	.37	−.27	.026
3rd choice, Patterns	−.01	.06	.496
Ring Toss	.83	.86	.500
Independence Training			
Self-reliance Training:			
No. of specific directions	.28	.31	.460
No. of nonspecific directions	.01	.33	.159
No. of pushing statements	−.02	.12	.348
Autonomy:			
No times "initiator"	−.248	−.728	.026
No. times "decider"	−.258	−.276	.480
No. acts in decision process	.086	−.149	.245
Sanctions			
No. of Warmth responses	−.39	.23	.030
No. of Rejection responses	.23	.70	.032

In time he comes to have the same expectations of himself. Learning to respond to such standards and expectations of high performance can be conceived of as learning a cognitive map of the world in which these standards and expectations are, so to speak, a relevant part of the terrain. It seems reasonable to assume that a strong set of standards is most likely to be learned when both parents agree on standards of excellence and have high aspirations and expectations for achievement. Where agreement between parents is lacking the boy is more likely to be confused and the probability of the standards becoming cues for affective arousal is lowered.

For the tasks performed in this study the optimal learning situation would

be one in which the parental aspirations and choices are in agreement or, in terms of the correlation analysis of parental pairs, one in which the parents play similar roles. The data show that for several achievement training measures (estimates and choices in Block Stacking and Patterns) the parents of high *n* Achievement boys tend to play more similar roles than do the parents of low *n* Achievement boys. That is, six of the seven correlations are positive and five are higher for the parents of the "highs" than for the parents of the "lows." The extremely high correlations for Total Choices, Ring Toss, is a result of the fact that each parent was asked to make a series of successive choices with both parents basing each choice on the previous position of the boy. There is one significant difference between the correlations or the parents of high and low *n* Achievement groups: the second choice in the Patterns task. This is interesting when we remember that the patterns chosen by the parents of the "highs" were not significantly more difficult than those chosen by the parents of the "lows." There is, however, greater agreement between the parents of high *n* Achievement boys as to how difficult the task should be, and in general this is more likely to be the case among parents of high than of low *n* Achievement boys.

The situation is reversed for the measures of self-reliance training, i.e., S, N, and P type statements. For these variables, the parents of low *n* Achievement boys tend to have the more similar roles. Thus, for all three variables the correlations are positive and are higher between parents of the "lows," although the differences between the two groups are not statistically significant. This finding lends additional credence to our hypothesis that self-reliance training, when not associated with high achievement training, is not a sufficient cause of *n* Achievement. The consistency with which both parents of low *n* Achievement boys put their sons on their own is especially significant in the light of the greater dissimilarity in their willingness to let the boy assert himself in the decision-making process. The correlations between fathers and mothers of low *n* Achievement boys for the three measures of autonomy are all negative and higher than the correlations for the parents of the "highs." For one measure, "number of times initiator," the difference between high or low groups is statistically significant at the .026 level. This greater dissimilarity in parental roles indicates a tendency for one person to dominate the decision-making process and suggests a greater loss of autonomy for low *n* Achievement boys. Where both parents try to affect the decision-making process the boy has some freedom of movement, and can, in the terminology of Simmel, even play the role of the "third who enjoys," but where one parent lays down the law the boy has fewer alternatives. It should be noted that in two of the measures of autonomy the parents of high *n* Achievement boys also play somewhat dissimilar roles, and that for the measure "number of times de-

cider" there is virtually no difference between "high" and "low" groups, which is probably a function of this aspect of the decision-making process where only one person wins.

The affect given out by the parents is considered to be the basic factor determining the child's affective arousal to standards of excellence. We believe that the child's sensitivity to parental affect is increased when he is faced with a situation where there is a difference between parents in the amount and kind of affect they display, or where a single parent is both warm and rejecting. Positive affect (Warmth), in particular, becomes more valued as it is both given and withheld; thus a parent who is both warm and rejecting will have more effect on the child than one who is uniformly warm. Hence, both dissimilarity between parents, or dissimilar affect from one parent, will result in a greater value placed upon warmth by the child. This notion is supported by a finding by Sears, Maccoby, Levin, *et al.* (12) who note that the child who has "high conscience" has a mother who is relatively warm and uses withdrawal of love fairly often; neither warmth nor withdrawal of love when used alone is effective in producing "high conscience."

This finding has particular relevance for the data on positive sanctions shown in Table 1 where it can be seen that the parents of the "highs" play far more dissimilar roles than the parents of the "lows"; the correlation between the fathers and mothers of high *n* Achievement boys for Warmth is —.39, as compared with .23 for the parents of low *n* Achievement boys. This difference between groups is significant at the .03 level. It should be remembered that both parents of the "highs" were above average for Warmth, although this difference was not significant for fathers, and that the mothers of high *n* Achievement boys were more warm and rejecting than the mothers of the "lows." We believe that this dissimilarity between parents and variations in the affect given out by the mother sensitizes the high *n* Achievement boy to the need for approval and increases his willingness to internalize the standards and expectations of his parents.

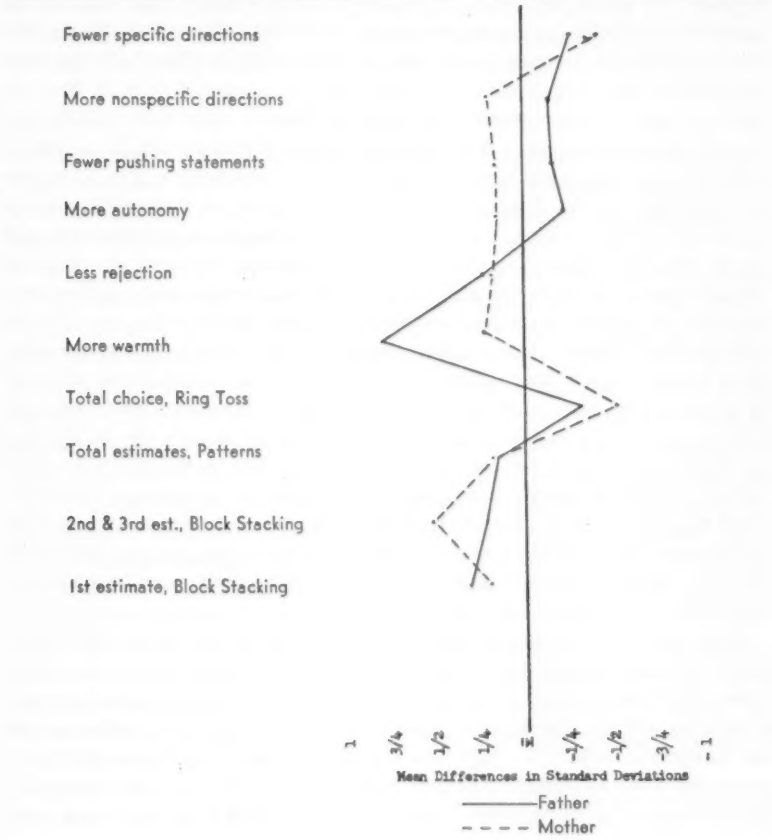
The role relationships of parents with respect to Rejection is unlike that found for Warmth. Correlations for this variable are positive for both groups, but the correlation is higher for the parents of the "lows" (.70) than for the parents of the "highs" (.23): the difference is significant at the .032 level. These data indicate that when the low *n* Achievement boy experiences rejection he is far more likely to receive it from both parents than is the high *n* Achievement boy. Negative affect, then, when not associated with compensating warmth from one or both of the parents, seems not likely to generate high need achievement.

Social Class and Child-Training Practices

The research design made it improbable that the class differences in training practices would be great, for the sample was selected in such a way as to make both classes equal with respect to the dependent variable. That is, the *n* Achievement scores of lower class boys were on the average equal to the scores of middle-class boys—a situation known not to be true for the more general universe of middle- and lower class boys. Hence, if the independent variables examined in this study are necessary and sufficient causes of *n*

CHART 4

Profiles for Social Class, Father and Mother, Using Mean Differences in Standard Deviations



Achievement, and the measures of these variables valid, there should be no significant differences in the training practices of the middle- and lower class parents *in this sample*. Significant class differences would be found only if these variables were noncausally related to *n* Achievement but associated with social class—a conclusion which had not been apparent when the original hypotheses were formulated.

The data shown in Chart 4 reveal class differences, but with one exception they are not significant; nor is it always the middle-class group whose scores are in the direction hypothesized as generating high *n* Achievement. The profiles shown in Chart 4 were derived in a manner similar to that described in the discussion of Chart 3. In this instance each profile is obtained by computing the mean difference between the scores of the fathers and mothers of middle- and lower class boys. A positive difference means that the average middle-class parent has the higher score; a negative difference indicates that the score for the average lower class parent is higher. The chart has been designed so that a high score is in the direction predicted as most likely to produce high *n* Achievement. The data in Chart 4 show that middle-class parents tend to emphasize achievement training; they have higher aspiration scores for the Block Stacking and Patterns tasks and tend to give their sons letters sooner in the Anagrams task. However, none of these differences is significant at the .05 level. Furthermore, the differences are somewhat affected by the boys performance; in every experiment but Ring Toss and Hatrack task (which none of the boys was able to solve) the middle-class boy's performance tended to be superior to that of his lower class peer. Middle-class boys, on the average, built higher towers, made patterns more quickly, and constructed more words. None of these differences, however, is significant at the .05 level. It would appear then when I.Q. and achievement motivation levels are controlled, social class membership is not a significant variable in determining performance, at least for the tasks included in this study. The one clear difference between middle- and lower class parents for a measure of achievement training is Total Choices, Ring Toss. In this case it is the lower class parent who tends to place the boy farthest away from the peg ($F=6.99$, $P<.01$), quite in variance with what we had anticipated, and for which no satisfactory explanation can be offered.

With respect to the independence training variables (S, N, P, and autonomy), it is the direction rather than the statistical significance of the differences which warrant their being reported. Middle-class mothers give somewhat more independence training than lower class mothers: they have higher scores for nonspecific directions, make fewer pushing remarks, and give their sons somewhat more autonomy; however, they also give more specific directions. The direction of all these scores is exactly the reverse of what was found to be

characteristic of the mothers of high *n* Achievement boys. The profile for fathers is almost point for point different. In this case it is the lower class father who gives fewer specific statements, more nonspecific statements, is less pushing, and gives the boy more autonomy. All these data are exactly contrary to what we had expected. It is important to note, however, that none of these differences is statistically significant at the one-in-ten level.

In the case of our two measures of sanctions, Warmth and Rejection, we found, as was expected, that the middle-class father and mother are warmer and less rejecting than lower class parents. The greatest difference is among fathers ($F=4.13, P<.06$).

These data, though sometimes in the direction predicted, indicate that the training practices of middle-class parents are not on the whole markedly different from those of lower class parents. However, it may be that by controlling *n* Achievement we in effect cancelled out for this sample any differences in training practices which might normally differentiate the social classes.

DISCUSSION AND SUMMARY

The question of how achievement training, independence training, and sanctions are related to achievement motivation may be rephrased by asking, How does the behavior of parents of boys with high *n* Achievement differ from the behavior of parents whose sons have low *n* Achievement?

To begin with, the observers' subjective impressions are that the parents of high *n* Achievement boys tend to be more competitive, show more involvement, and seem to take more pleasure in the problem-solving experiments. They appear to be more interested and concerned with their son's performance; they tend to give him more things to manipulate rather than fewer; on the average they put out more affective acts. More objective data show that the parents of a boy with high *n* Achievement tend to have higher aspirations for him to do well at any given task, and they seem to have a higher regard for his competence at problem solving. They set up standards of excellence for the boy even when none is given, or if a standard is given will expect him to do "better than average." As he progresses they tend to react to his performance with warmth and approval, or, in the case of the mothers especially, with disapproval if he performs poorly.

It seems clear that achievement training contributes more to the development of *n* Achievement than does independence training. Indeed, the role of independence training in generating achievement motivation can only be understood in the context of what appears to be a division of labor between the fathers and mothers of high *n* Achievement boys.

Fathers and mothers both provide achievement training and independence training, but the fathers seem to contribute much more to the latter than do

the mothers. Fathers tend to let their sons develop some self-reliance by giving hints (N) rather than always telling "how to do it" (S). They are less likely to push (P) and more likely to give the boy a greater degree of autonomy in making his own decisions. Fathers of high *n* Achievement boys often appear to be competent men who are willing to take a back seat while their sons are performing. They tend to beckon from ahead rather than push from behind.

The mothers of boys with high achievement motivation tend to stress achievement training rather than independence training. In fact, they are likely to be more dominant and to expect less self-reliance than the mothers of boys with low *n* Achievement. But their aspirations for their sons are higher and their concern over success greater. Thus, they expect the boys to build higher towers and place them farther away from the peg in the Ring Toss experiment. As a boy works his mother tends to become emotionally involved. Not only is she more likely to reward him with approval (Warmth) but also to punish him with hostility (Rejection). *In a way, it is this factor of involvement that most clearly sets the mothers of hign* n *Achievement boys apart from the mothers of low* n *Achievement boys:* the former score higher on every variable, expect specific directions. And although these mothers are likely to give their sons more option as to exactly (fewer Specifics) what to do, they give them less option about doing something and doing it well. Observers report that the mothers of high *n* Achievement boys tend to be striving, competent persons. Apparently they expect their sons to be the same.

The different emphasis which the fathers and mothers of high *n* Achievement boys place upon achievement and independence training suggests that the training practices of father and mother affect the boy in different ways. Apparently, the boy can take and perhaps needs achievement training from both parents, but the effects of independence training and sanctions, in particular Autonomy and Rejection, are different depending upon whether they come from the father or mother. In order for high *n* Achievement to develop, the boy appears to need more autonomy from his father than from his mother. The father who gives the boy a relatively high degree of autonomy provides him with an opportunity to compete on his own ground, to test his skill, and to gain a sense of confidence in his own competence. The dominating father may crush his son (and in so doing destroys the boy's achievement motive), perhaps because he views the boy as a competitor and is viewed as such by his son. On the other hand, the mother who dominates the decision-making process does not seem to have the same affect on the boy, possibly because she is perceived as *imposing her standards* on the boy, while a dominating father is perceived as *imposing himself* on the son. It may be that the mother-son relations are typically more secure than those between father and son, so that the boy is better able to accept higher levels of dominance and rejection

from his mother than his father without adverse affect on his need to achieve. Relatively rejecting, dominating fathers, particularly those with less than average warmth—as tended to be the case with the fathers of low *n* Achievement boys—seem to be a threat to the boy and a deterent to the development of *n* Achievement. On the other hand, above-average dominance and rejection, coupled with above-average warmth, as tends to be the case with mothers of high *n* Achievement boys, appear to be a spur to achievement motivation. It will be remembered that the fathers of high *n* Achievement boys are on the average less Rejecting, less Pushing, and less Dominant—all of which points to their general hands-off policy.

It is unlikely that these variables operate separately, but the way in which they interact in the development of achievement motivation is not clear. Possibly the variables interact in a manner which produces cyclical effects roughly approximating the interaction that characterized the experimental task situations of this study. The cycle begins with the parents imposing standards of excellence upon a task and setting a high goal for the boy to achieve (e.g., Ring Toss, estimates and choices in Block Stacking and Patterns). As the boy engages in the task, they reinforce acceptable behavior by expressions of warmth (both parents) or by evidences of disapproval (primarily mother). The boy's performance improves, in part because of previous experience and in part because of the greater concern shown by his parents and expressed through affective reaction to his performance and greater attention to his training. With improved performance, the parents grant the boy greater autonomy and interfere less with his performance (primarily father). Goals are then reset at a higher level and the cycle continues.

Manuscript received: July 28, 1958
Revised manuscript received: February 27, 1959

Bernard C. Rosen
Department of Sociology
University of Connecticut
Storrs, Connecticut

REFERENCES

1. Bales, R. F., *Interaction Process Analysis,* Cambridge, Mass.: Addison-Wesley, 1951.
2. Child, I. L., T. Storm, and J. Veroff, "Achievement Themes in Folk Tales Related to Socialization Practice," in J. W. Atkinson, *Motives in Fantasy, Action and Society,* Princeton, N. J.: Van Nostrand, 1958.
3. Erickson, M. C., "Social Status and Child-rearing Practices," in T. Newcomb and E. Hartley, *Readings in Social Psychology,* New York: Holt, 1947.
4. Havighurst, R. J., and A. Davis, "Social Class Differences in Child-rearing," *American Sociological Review,* 1955, 20, 438–442.
5. Hollingshead, A., and F. C. Redlick, "Social Stratification and Psychiatric Disorders," *American Sociological Review,* 1953, 18, 163–169.

6. McClelland, D. C., and G. A. Friedman, "A Cross-Cultural Study of the Relationship between Child-Training Practices and Achievement Motivation, appearing in Folk Tales," in G. E. Swanson, T. M. Newcomb, and E. L. Hartley, eds., *Readings in Social Psychology*, New York: Holt, 1952.
7. McClelland, D. C., J. W. Atkinson, R. Clark, and E. Lowell, *The Achievement Motive*, New York: Appleton-Century-Crofts, 1953.
8. McClelland, D. C., A. Rindlisbacher, and R. deCharms, "Religious and Other Sources of Parental Attitudes toward Independence Training," in D. C. McClelland, ed., *Studies in Motivation*, New York: Appleton-Century-Crofts, 1955.
9. Rosen, B. C., "The Achievement Syndrome: a Psychocultural Dimension of Social Stratification," *American Sociological Review*, 1956, 21, 203–211.
10. Rosen, B. C., "Race, Ethnicity, and the Achievement Syndrome," *American Sociological Review*, 1959, 24, 47–60.
11. Sakoda, J. M., "Directions for a Multiple-Group Method of Factor Analysis," mimeographed paper, University of Connecticut, June, 1955.
12. Sears, R. R., E. E. Maccoby, and H. Levin in collaboration with E. L. Lowell, P. S. Sears, and J. W. M. Whiting, *Patterns of Child Rearing*, Evanston, Ill.: Row, Peterson, 1957.
13. Strodtbeck, F. L., "Family Interaction, Values, and Achievement," in D. C. McClelland, A. L. Baldwin, U. Bronfenbrenner, and F. L. Strodtbeck, *Talent and Society*, Princeton, N. J.: Van Nostrand, 1958.
14. Tryon, R. C., *Cluster Analysis*, Ann Arbor, Mich.: Edwards Brothers, 1939.
15. Winterbottom, M. R., "The Relation of Need for Achievement to Learning Experiences in Independence and Mastery," in J. W. Atkinson, *Motives in Fantasy, Action, and Society*, Princeton, N. J.: Van Nostrand, 1958.

Cognitive Consistency, Response Reinforcement, and Attitude Change [1]

WILLIAM A. SCOTT, *University of Colorado*

Attitudes have been conceptualized both as dispositions to behave (2; 3; 14, pp. 228–229) and as properties of the person's cognitive structure which summarize his cognitive, affective, emotional, and motivational orientations toward the concept of some object or event (6, pp. 151–152; 7). Just what properties are attributed to attitudes will depend, in part, on whether one's conception of them focuses on their behavioral or on their cognitive aspects. Within either framework it is customary to recognize the properties of *direction* and *strength* (6, p. 156; 14, pp. 228–229). However, structural characteristics of attitudes —such as *precision, specificity, differentiation, and hierarchic integration*— are more frequently elaborated by the cognitive theorists (e.g., 5; 6, pp. 158–167) than by behaviorists of the S-R learning school.

The feature common to such structural attributes is that they refer to relationships which attitudes bear to other cognitive elements. Despite their profuseness in theory, they rarely appear in research designs. Empirical studies of attitudes have dealt, for the most part, with the simple properties of *direction* and *strength,* rather than with more elaborate relational characteristics. Perhaps the rarity of the latter in attitude research is largely due to the difficulty of operationalizing concepts so subtle and complex as these frequently are. Certainly such a difficulty was encountered in the present study.

The structural property of attitudes investigated here was that which has been referred to elsewhere (10) as *cognitive consistency*. This refers to the degree to which the attitude is "embedded" within an attitude structure consisting of values and expectancies related to the event which is its focus. Thus, the concept is akin to those of *unity* and *hierarchic integration* proposed by Krech and Crutchfield (6, p. 162) and by French (5).

A cognitively consistent attitude is one which follows logically from the person's view of an event in relation to his goals. More precisely, an attitude is said to be cognitively consistent to the degree that (a) the values on which

[1] The research reported in this article was supported by a grant from the National Institute of Mental Health. It grew out of a pilot study (12) conducted at the Behavioral Science Conference, University of New Mexico, in the summer of 1957, under the auspices of the U. S. Air Force Office of Scientific Research, Contract No. 49 (638)-33. I am indebted to Mrs. Brenda Dickey for technical assistance in the conduct of studies and analysis of the data.

it rests are not inherently contradictory, (b) the focal event is clearly perceived in relation to the values, and (c) the valence of the attitude (both direction and strength) is congruent with the individual's perception of the object-goal relationships.

This last aspect of cognitive consistency is the critical one; it may be clarified by an illustration: If a person values free enterprise, and sees a corporation's excess profits tax as antithetical to the free enterprise system, then he "should" be against the tax—providing other goals and expectancies do not counteract this disposition. If he is actually against it, one would term his attitude toward the tax cognitively consistent (given this information only). However, if he were in favor of the tax, his attitude would be inconsistent with the value of free enterprise and the expected relevance of the event (corporation tax) to the value.

In the normal situation, any particular event is likely to be seen in relation to a number of values; some of these expectancies may be favorable (the value is enhanced by the event), and some unfavorable (the value is threatened by the event). Under such circumstances a cognitively consistent attitude is one which corresponds to the balance of favorable and unfavorable consequences that the event is seen to portend. Arriving at this balance involves a complex integrative process in which the various relevant values are assigned different weights and the event is judged differentially relevant to each of them.

Some theorists (e.g., 1) appear to assume that all attitudes are cognitively consistent, and the only thing which keeps them from appearing so is the researcher's failure to assess all the relevant values and expectancies and to weight them properly in the total. Indeed, there is rather impressive empirical evidence (8) that if enough cognitive components are assessed, with sufficient precision, they may permit rather accurate prediction of the attitude which depends on them.

It is possible, however, to maintain a somewhat divergent view: Attitudes need not be cognitively consistent, unless the requirements of external reality (physical or social) impose such consistency on them. It is quite possible for one to hold an attitude for no consciously defensible reason whatsoever, or to hold an attitude in spite of contradictory reasons which are recognized by the person himself as valid. Such an attitude would be regarded by an investigator as cognitively inconsistent, regardless of the precision and searchingness of his inquiry.

In order for an attitude to be cognitively consistent by the above definition, it is not necessary for it to be rational—i.e., it need not correspond with the "true" consequences of the event for one's values, as seen by some omniscient outside observer. It is only necessary for the person to perceive the event in relation to his goals in such a way that the attitude follows as a logical con-

sequence. Also, there is no need for a cognitively consistent attitude to be univalent. If one perceives an event as boding both favorably and unfavorably, in about equal degrees, then a cognitively consistent attitude would be a neutral, or ambivalent, one.

Festinger (4) has proposed the term *cognitive dissonance* to describe (among other things) attitudes which are not univalent and suggests that such a condition will give rise to cognitive strains to reduce the dissonance. If commitment to the attitude is irrevocable, reduction in dissonance might typically be achieved by reducing the strengths of the competing values or by altering one's ideas about the relevance of the event to them.

The present view implies no such antipathy for cognitive dissonance. Ambivalent attitudes may be normal, stable components of many people's cognitive structures. They may even be quite cognitively consistent. Moreover, cognitive consistency itself is not even a requisite for a quasi-stable state of the total cognitive structure. Only when pressures for such consistency arise from within the person himself or from his environment do the demands for adaptation require a more consistent attitude structure. It is quite possible for one not confronted with the need to adjust to a consistently ordered environment, or with a need to defend his opinions against opponents who demand to know their rational basis, to maintain quite inconsistent attitudes. He may simply isolate the particular attitude structure from the rest of his cognitive systems, or he may fail to impose on himself any requirements for rationality in that area.

Given that any particular attitude structure may be cognitively consistent or not, one might suggest that the degree of consistency would have important implications for the stability of the attitude. If an attitude structure be regarded as a system of interrelated cognitive elements (10), then the closeness of their interrelationship should help determine the stability of the system. Thus, a cognitively consistent attitude should be more resistent to change than an inconsistent one, because it forms a more Prägnant system with the other cognitive elements, values, and expectancies. (The above principle should be read *ceteris paribus,* since an attitude might well be in systemic relation with other cognitive or personality elements besides values and expectancies, and thus maintain relative stability regardless of its degree of cognitive consistency, in the present sense.)

An experiment was designed to test whether or not cognitive consistency was an important determinant of attitude stability. From previous studies (9, 13) had evolved a rather dependable procedure for inducing attitude change, namely by having Ss present, orally, opinions divergent from their own attitudes and then rewarding their verbal behaviors. Quite consistently these experiments had shown that the "winners" of debates (arbitrarily deter-

mined) tended to change their attitudes in the direction of their presentations more than did "losers" or control groups of nondebaters. Therefore, an equivalent method of response reinforcement was used to create pressures toward attitude change. Attitudes of Ss prior to the induced verbal behavior were assessed, by a fairly crude method, as either cognitively consistent or inconsistent. It was predicted that, under the experimental conditions, the cognitively consistent Ss would be more likely to maintain their original attitudes than the inconsistent ones.

METHOD

Assessment of Cognitive Consistency

The concept of cognitive consistency can be operationalized in a variety of ways, depending on what related cognitive elements one chooses to assess and on how one chooses to define a "consistent" relationship among them. The particular operational definition employed in this study was adapted from a procedure developed by Rosenberg (8). An attitude was deemed to be cognitively consistent to the extent that its direction and strength could be predicted from a knowledge of S's values and expectancies in relation to the object of the attitude.

The method used by Rosenberg involved independent assessment of three types of cognitive elements: An attitude toward X, S's values or goals, and the degree to which X was perceived as relevant to the values, in a positive or negative fashion. A simple multiplicative model was employed to predict S's attitude (A') from the sum of the products of values (Vs) and perceived relevances (Rs). Thus: $A' = \sum_{i=1}^{k} V_i R_i$, where k represents the number of values assessed by the instrument. Rosenberg obtained very high correlations between predicted attitudes and actual attitudes as independently assessed. In previous and current studies by the writer (12) such high correlations have never been obtained. This discrepancy may be due in part to less precise assessment procedures and to the focus on attitudes of less central concern to Ss. In any case, the failure of perfect correlation between predicted and actual attitudes permits one to distinguish varying degrees of cognitive consistency. Whereas nearly all of Rosenberg's Ss would here be regarded as cognitively consistent, a much smaller proportion of the present experimental group would be so judged.

The attitudes of Ss toward Greek organizations and toward campus activities were measured by two closed questions:

All things considered, do you think fraternities and sororities are a help to students at CU or a hindrance?

.... a. A great help
.... b. More of a help than a hindrance
.... c. Both a help and a hindrance, depending on the individual
.... d. More of a hindrance than a help
.... e. A great hindrance

Which do you think the campus needs more of:

.... a. More intellectual activity, such as debates, exhibits and concerts; and less friendly, interpersonal activity, such as dances, hobby clubs and outings
.... b. More intellectual activity, keeping the present amount of friendly, interpersonal activity
.... c. No change necessary
.... d. More friendly, interpersonal activity, keeping the present amount of intellectual activity
.... e. More friendly, interpersonal activity; and less intellectual activity

Selected personal values were assessed by a series of items referring to eight different types of "ideal relations" among people. (A *personal value,* or *moral ideal* is defined (14, p. 226) as a concept of an ideal relation among actors, or between an actor and his environment, which serves as a standard for judging the "goodness" or "badness," the "rightness" or "wrongness," of any particular observed relation.) These eight values were selected because they were among those most frequently mentioned by students when asked to describe ideal traits in people (11) they were also deemed, a priori, to be potentially relevant to Ss' evaluations of Greek organizations and campus activities.

Here are some statements that might be regarded by some people as "ideal traits" to be developed in themselves and others. Check how you feel about each of them in the appropriate space. (Each statement was followed by the five alternatives indicated for the first value.)

1. Being charming, popular, well-mannered, and getting along with all kinds of people.
   .... a. Admire very much
   .... b. Admire somewhat
   .... c. Neither admire nor dislike; depends on circumstances
   .... d. Dislike somewhat
   .... e. Dislike very much
2. Being inventive, creative, and always thinking of different ways of doing things.
3. Having strong intellectual and cultural interests; trying to learn a great deal about things, even though the knowledge may not be useful.
4. Studying a great deal and working hard to get good grades.
5. Being mostly concerned about other people, doing good for them and trying to make them happy, even if it is against one's own interests.
6. Always telling the truth and being completely honest; never cheating or lying, even though these might make for an easier relationship with others.
7. Having strong leadership qualities, being respected by others, and gaining recognition for one's achievements.

8. Being independent, outspoken, freethinking, and unhampered by the bounds of tradition or social restraint.

Also administered was a measure of *expectancies* or *perceived relevance* of object X to the values, as follows:

In deciding whether fraternities and sororities are a help or a hindrance to students at CU, how do you see the following ideals as relevant? Write down a number from the following scale in the space beside each ideal.

SCALE

1	2	3	4	5
Fraternities and sororities are very favorable to this ideal	Fraternities and sororities are somewhat favorable to this ideal	Fraternities and sororities are neither for nor against this ideal	Fraternities and sororities are somewhat against this ideal	Fraternities and sororities are very much against this ideal

.... a. Being charming, popular, well-mannered, and getting along with all kinds of people.
.... b. Being inventive, creative and always thinking of different ways of doing things.
.... c. Having strong intellectual and cultural interests; trying to learn a great deal about things, even though the knowledge may not be useful.
.... d. Studying a great deal and working hard to get good grades.
.... e. Being mostly concerned about other people, doing good for them, and trying to make them happy, even if it is against one's own interests.
.... f. Always telling the truth and being completely honest; never cheating or lying, even though these might make for an easier relationship with others.
.... g. Having strong leadership qualities, being respected by others, and gaining recognition for one's achievements.
.... h. Being independent, outspoken, freethinking, and unhampered by the bounds of tradition or social restraint.

In deciding whether the campus needs more intellectual activity or more friendly, interpersonal activity, how do you see the following ideals as relevant? Write down a number from the following scale in the space beside each ideal.

SCALE

1	2	3	4	5
This ideal strongly favors more intellectual activity	This ideal somewhat favors more intellectual activity	This ideal is irrelevant to the question	This ideal somewhat favors more friendly interpersonal activity	This ideal strongly favors more friendly interpersonal activity

(Same statements, a–h as for preceding question.)

Each value strength (V_i) and each value relevance (R_i) was checked on a scale which was subsequently converted to scores between -2 and $+2$. The

contribution of any particular value to S's attitude toward X could be arbitrarily computed as the product of the value strength and the corresponding relevance (V_iR_i). The total contribution of all measured values to the attitude could be arrived at by summing the products of value strengths and relevances over all eight values. Thus, a predicted attitude, A', arrived at by this procedure, would be:

$$A' = \sum_{i=1}^{k} V_i R_i$$

S's actual attitude, A, had been assessed by the questions quoted above. A' would presumably provide a predictor of A to the extent that (1) all relevant values were assessed, (2) S's attitude was wholly determined by the values and perceived relevances, and (3) the methods of attitude determination were adequately represented by the ΣVR model utilized here. Requirements (1) and (3) were taken as assumptions (though they were probably not met) and failure of prediction of A from A' was attributed to cognitive inconsistency, or departure from requirement (2). The degree of cognitive inconsistency in a particular S was thus inferred from the discrepancy between A and A'.

For a particular X (Greek organizations or campus activity), the ΣVR distribution was divided into five segments, corresponding in frequencies to the distribution on the five-category attitude continuum. If a particular S's ΣVR score placed him in a group corresponding to his actual attitude score, he was regarded as cognitively consistent; if not, he was called cognitively inconsistent. (It is evident that, with a five-category attitude scale, there could be a maximum of four degrees of inconsistency; in the present study, actual attitude scores were rarely more than one category discrepant from their corresponding ΣVR score categories, so all inconsistent Ss were grouped together, regardless of the degree of inconsistency.)

This procedure provided a method of labeling Ss, crudely, as cognitively consistent or inconsistent with regard to a particular attitude, on a pretest which was administered about a week prior to the attitude change experiment.

Experimental Procedure

The experimental arrangements were aimed at inducing Ss to present opinions different from their own, and then rewarding them for it. Accordingly, they were asked to prepare to debate one of the issues, taking a side opposite to that which they had indicated as their own on the pretest. The request was made about a week before the time they were scheduled to appear. Ss were told that some people were going to present their own true views, while others would take an opposite position, since the purpose of the experiment was to see how well the participants could judge each others' true attitudes from what

the person said in the debate. Both cognitively consistent and inconsistent Ss were selected to participate, in a random sequence which was unknown to E.

Ss appeared individually in a small research room, and were introduced to "another subject" (a confederate). These two persons sat at one end of a long table, and three judges (E and two graduate students) at the other end. The confederate was used simply to maintain the pretext of the study. The intent was to "reward" all Ss for their presentations, so that their relative degrees of attitude stability could be studied.

The first mechanism of reward chosen was the one which had proved successful in previous experiments—i.e., to have S present an initial argument, followed by the confederate's presentations (gauged to an appropriate level of finesse) and rebuttals in reverse order, then to tell S that he had won the debate and offer him a $2 prize. It became apparent, however, in the first several debates, that winning either a monetary reward or the judges' approval was not an effective reinforcing device. Ss rarely changed their attitudes. This may have been due in part to the kind of measure used and to the fact that they had already checked their answers on the same scale three or more times previously.[2] But it was also considered likely that an inappropriate reward had been chosen. In contrast with the recruitment procedure for previous experiments of this type, these Ss had not been asked to volunteer for debates. Initially, they had only been asked to participate in some discussions which would help E to "learn something about what kinds of people hold what kinds of attitudes." During the two one-hour discussion sessions preceding the debate, E's and the graduate assistants' demeanors had been informal and friendly. Therefore, it seemed quite likely that these particular Ss, at this stage in the study, were not particularly motivated to win a competitive debate, so the intended "reward" was not a reward at all.

In an effort to give the largest possible number of Ss a "rewarding experience," it was decided to abandon the formal debate procedure in favor of an informal discussion between S and the confederate. In the presence of the three judges S was first asked to present his side of the issue (still opposite

[2] Other objectives of the study required that Ss participate in free discussions on these issues for two one-hour sessions prior to the present experiment. Their attitudes, values, and expectancies had been assessed, by identical instruments, two or more times prior to the pretest assessment reported here. It is quite certain, therefore, that many of them developed response sets or became "attached" to the numbers they had previously marked on the scale. This may account for the relatively little change in attitude induced by the experimental conditions, in comparison with the amount achieved in previous studies (9, 13). Furthermore, in the previous experiments, attitudes had been assessed from Ss' free responses to open questions. This also presumably gave them greater opportunity to change than did the present closed-question technique, since one can forget words more readily than numbers.

from his own true position), followed by appropriate impromptu remarks from the confederate, supporting the other side. S was then given an opportunity to reply. The total discussion lasted about ten minutes, during which the three judges took notes and showed a sympathetic interest in what was said. Then E thanked both participants for their "very interesting ideas," in an effort to reward S by making him feel that he had performed well, and conducted both S and the confederate to separate rooms adjoining the experimental room. There S took an attitude posttest which also included questions about four irrelevant issues and about S's estimate of his opponent's true opinion concerning the topic just discussed.

Under such an arrangement only two Ss "broke out of role," by overtly professing a position contrary to that which they were instructed to assume. Data on these two are included in the following analysis. Altogether 15 of the Ss were treated as debaters, 23 as discussants.

RESULTS

Since all 38 Ss were presumably exposed to rewarding conditions, the first hypothesis, following the theory of response reinforcement, was that, for the group as a whole, there would be a mean shift in attitude in the direction of the assumed position. Table 1 shows that this hypothesis was confirmed, though the mean shift was not nearly so great as in the previous experiments. (Since the expected direction of attitude change in some Ss was toward the mean, part of the total shift in attitude might be accounted for by a regression effect. In order to control for this artifact, shifts in attitude toward the issue discussed were compared with shifts in some other attitude on which S had indicated a comparable pretest position. Since the mean change in the discussed attitude is significantly greater than the mean change in the nondiscussed attitudes, it may be safely concluded that the presumably reinforcing conditions of the experiment tended to produce the desired effect.)

TABLE 1

Mean Attitude Changes Following Reward

(A positive sign indicates a mean change in the direction of S's presentation; or, for nondiscussed issues, a mean change opposite to S's original position. When Ss' initial attitudes toward nondiscussed issues were neutral, the directions of change were assigned alternatingly positive and negative signs.)

Issue	N	Mean Change	S.D. of Change	Difference in Mean Changes
Discussed	38	+0.45	0.82	t (correlated measures) = 2.29
Not discussed	38	+0.11	0.51	$p < .05$

The major concern in this study, however, was with the differential effects of response reinforcement on cognitively consistent and inconsistent Ss. If cognitive consistency exerts a stabilizing influence on attitudes, then consistent Ss should tend to show less movement under the experimental conditions than inconsistent Ss. The figures reported in Table 2 indicate that this hypothesis was confirmed: More cognitively inconsistent than consistent Ss changed their attitudes in one direction or the other. However, since some of them moved away from the side they had been forced to present, the mean change for the two groups was about the same. It is thus clear that cognitive incon-

TABLE 2

Frequencies of Attitude Change among Cognitively Consistent and Inconsistent Ss

Group of Subjects	N	Changed (in either direction)	Did not change	
Cognitively consistent	21	6	15	$\chi^2 = 6.83$
Cognitively inconsistent	17	12	5	$p < .01$

sistency did not permit prediction of the direction of attitude change under these conditions, but rather prediction of whether or not change (either toward or away from the rewarded position) occurred at all. Cognitively consistent Ss seemed more resistant to change in either direction.

SUMMARY

This experiment was designed to test the hypothesis that Ss whose attitudes are cognitively consistent are more resistant to attitude change under conditions of response reinforcement than are Ss with cognitively inconsistent attitudes. Cognitive consistency was here defined as congruity of an attitude with one's perception of an event in relation to his values. Attitudes of Ss toward several different issues were assessed by questionnaires, which also asked them to indicate the degree to which they admired eight different "ideal traits" and also the degree to which the object of each attitude would help or hurt each of the eight ideals. On the basis of these latter two sets of replies Ss' attitudes were "predicted," using a simple sum-of-products formula. Ss whose predicted attitudes corresponded with their actual attitudes were called cognitively consistent, and the rest were called inconsistent.

In random order consistent and inconsistent Ss were induced to present arguments on one of two issues, always taking a side different from that which they had indicated as their own on the pretest. Their verbal behaviors were "reinforced" with praise from the experimenter in an attempt to alter their attitudes in the direction of the presentation. While the group as a whole

showed a significant (though small) attitude shift in the direction of the reinforced behaviors, cognitively consistent Ss were much more likely to retain their initial attitudes than the inconsistent ones. This result was interpreted as indicating that consonance of an attitude with other cognitive elements, such as values and expectancies, serves to stabilize it and increase its resistance to change under externally imposed pressures.

Manuscript received: July 28, 1958
Revised manuscript received: January 12, 1959
William A. Scott
Department of Psychology
University of Colorado
Boulder, Colorado

REFERENCES

1. Abelson, R. P., and M. J. Rosenberg, "Symbolic Psychologic: A Model of Attitudinal Cognition," *Behavioral Science*, 1958, 3, 1–13.
2. Allport, G. W., "Attitudes," in C. Murchison, ed., *Handbook of Social Psychology*, Worcester, Mass.: Clark University Press, 1935.
3. Doob, L. W., "The Behavior of Attitudes," *Psychological Review*, 1947, 54, 135–156.
4. Festinger, L., "*A Theory of Cognitive Dissonance*," Evanston, Ill.: Row, Peterson, 1956.
5. French, V., "The Structure of Sentiments," *Journal of Personality*, 1947, 15, 247–282; 1948, 16, 78–100, 209–244.
6. Krech, D., and R. S. Crutchfield, *Theory and Problems of Social Psychology*, New York: McGraw-Hill, 1948.
7. Peak, H. "Attitude and Motivation," in M. R. Jones, ed., *Nebraska Symposium on Motivation: 1955*, Lincoln: University of Nebraska Press, 1955.
8. Rosenberg, M. J., "Cognitive Structure and Attitudinal Affect," *Journal of Abnormal and Social Psychology*, 1956, 53, 367–372.
9. Scott, W. A., "Attitude Change through Reward of Verbal Behavior," *Journal of Abnormal and Social Psychology*, 1957, 55, 72–75.
10. Scott, W. A., "Rationality and Non-rationality of International Attitudes," *Conflict Resolution*, 1958, 2, 8–16.
11. Scott, W. A., "Empirical Assessment of Values and Ideologies," *American Sociological Review*, 1959, 24, 299–310.
12. Scott, W. A., "Personal Values and Group Interaction," In United States Air Force, Office of Scientific Research, eds., *Decisions, Values, and Groups*, New York: Pergamon Press (in press).
13. Scott, W. A., "Attitude Change by Response Reinforcement: Replication and Extension," *Sociometry* (in press).
14. Scott, W. A., and S. B. Withey, *The United States and the United Nations: The Public View*, New York: Manhattan Publishing Company, 1958.

Human Interaction and Interpersonal Perception [1]

IVAN D. STEINER, *University of Illinois*

During the past five years there has been a great proliferation of journal articles linking human interaction with interpersonal perception. This development is not really very surprising because everyday experience suggests that the way we react to our fellow men depends in no small measure upon how we perceive and interpret their motives, preferences, and intentions.

It is the *content* of these many articles which is surprising. The case for saying that interpersonal perceptions influence human interaction has seemed so obvious and straightforward that one is not prepared to find it denied by negative or contradictory findings. Of course there have been many positive findings, too, but the negative ones have been frequent enough to suggest that something has been wrong. The purpose of this paper is to discuss some of the complexities which haunt this area of research and which may be responsible for some of the discouraging findings that have appeared in the literature. Many of the most troublesome methodological problems have already been discussed by other writers (1, 2, 4, 5, 7) and need not be reviewed here. Instead, this paper will focus upon problems which have received little or no attention in the literature.

ACCURACY OF INTERPERSONAL PERCEPTION

We can approach one of these complexities by considering a very popular hypothesis: if an individual accurately perceives and understands his associates' motives, preferences, and intentions, he can "get along" with those associates better than would be the case if his perceptions and interpretations were not so accurate. Accurate perceptions should permit the individual to gear his own behaviors into those of his associates so that a productive and harmonious dyad emerges. Everyday experience says this proposition is true; the trouble is that empirical research has often failed to confirm it. Individuals with high scores on accuracy of social perception are not always popular members of their groups, and when such people are formal leaders of work groups, their groups are not necessarily more effective than are groups led by persons whose accuracy scores are lower. Where does the trouble lie?

[1] This is a revised version of a paper presented as part of a symposium before the American Psychological Association, August, 1958.

One answer to this question has been provided by Taft (7) and by Gage and Cronbach (5) who have shown pretty clearly that our measurements of accuracy of interpersonal perception have sometimes been unreliable. Indeed, these authors have made their case so well that some investigators have indicated a willingness to give up the entire area of accuracy of interpersonal perception as one which is methodologically insoluble. This would be a serious mistake; one can grant everything that has been said about the difficulties involved in measuring accuracy and still work effectively with that variable. Instead of measuring accuracy, you manipulate it—and bear in mind the fact that one can often manipulate a variable when he cannot measure it very accurately. The cave men were manipulating temperature through the use of fire several thousand years before accurate thermometers were invented, and they certainly must have learned a good deal about the consequences of their manipulations.

Now it happens that those of us who say we are working with interpersonal perceptions have not often chosen to manipulate such perceptions. Instead we want to take people as they come and measure their perceptions without tampering with them. When we try to do this we possibly *are* attempting to accomplish more than is possible at the present stage of research sophistication. But researchers who work with interpersonal perceptions *without saying so* are often much more willing to manipulate their variables. One may cite as examples the numerous studies done by Festinger and his associates in which subjects are fed bogus or truthful information about one another, and their subsequent interpersonal behaviors are observed. These are not ordinarily called studies of interpersonal perception; instead they are said to be studies of cohesion, or power differentials, or group standards. But it is apparent that the experimental manipulations involve altering people's impressions of one another, and that these manipulations have often had marked effects on the human interaction which ensued.

In addition to the measurement problems, there are other reasons why research findings have often been disappointing. We have sometimes investigated the accuracy with which individuals perceive qualities in others which have no apparent relevance to the activities of the interacting individuals. This is very questionable procedure, for it has never been demonstrated that accuracy of social perception is a highly generalized phenomenon. Unless such generality can be demonstrated it is probably unreasonable to expect that individuals who accurately understand one another's food preferences or recreational interests are, for that reason, better able to function as a bomber crew or work group.

Occasionally we have also forgotten that perceptual accuracy is a tool which can be put to more than one use. Depending upon the motives of the persons

concerned, it can permit either effective cooperation *or* effective competition; it can also lead an individual to walk away and leave others to shift for themselves because their goals are accurately perceived to contradict his own. Without knowledge of people's motives it cannot be predicted that accurate social perceptions will generate harmonious or productive interpersonal behaviors.

Finally, in our eagerness to link accurate social perceptions with effective or harmonious interpersonal relations, we have sometimes overlooked the fact that a large percentage of all interpersonal behaviors occur within the limitations imposed by a role system. Such limitations often prevent people from treating one another as unique individuals, and require instead that each respond to the other as though he were a standard member of a category. Under such circumstances, the accuracy with which individuals perceive one another's idiosyncratic preferences or intentions is largely irrelevant, and accuracy of interpersonal perception can hardly be expected to generate harmonious or productive interpersonal relations.

Assumed Similarity

Some of the other complexities involved in research into interpersonal perceptions are most sharply revealed by studies of assumed similarity. This type of research is not concerned with *actual* similarity between people, but rather with similarity which is *assumed* by the individuals who interact with one another. Assumed similarity scores are generally obtained by having individuals rate two or more people, including themselves perhaps, on a number of continua. A D score or some other statistic is then used to summarize the magnitude of the discrepancies between the two or more sets of ratings.

This procedure is very attractive: it is easy to use; it seems to avoid some of the methodological difficulties inherent in accuracy scores; and it produces data which correlate with human interaction, or the products of human interaction. There can be no doubt that assumed similarity scores deserve our close attention.

The very term "assumed similarity" points to one of the difficulties in the use of such measurements. We are concerned with what the individual assumes to be true about himself and other people, regardless of whether or not his assumptions are correct. At least that is what we say we are concerned with. But it is often very difficult to interpret our findings without postulating that our assumed similarity scores tell us something about the individual's tendency to minimize or maximize similarities in an autistic manner. In short, our after-the-fact psychologizing usually leads us to think in terms of *unwarranted* assumed similarity or dissimilarity. Unfortunately, this is not what we have measured, and it is difficult to see how unwarranted assumed similarity can

be measured until someone invents a way of measuring accuracy of interpersonal perception. After all, one cannot say that a subject assumes more or less similarity than is warranted unless one has a dependable criterion of what is warranted—that is, of what constitutes accurate interpersonal perception. So some of the methodological problems which beset us when we study accuracy of interpersonal perception are still present when we shift to the study of assumed similarity.

An example will be helpful. Fiedler (3) has done some intriguing research on the relationships between assumed similarity scores of leaders and the effectiveness of their groups. Leaders of task-oriented groups are asked to think of someone with whom they have been able to work very well, and of someone with whom they have had great difficulty working. The leader rates each of these two persons on a series of continua, and a D score is derived to indicate the discrepancy between these two sets of ratings. Fiedler calls this the ASo score, or "Assumed Similarity between Opposites" score. The really interesting thing about this score is the fact that it correlates substantially with group effectiveness. Under at least certain specified conditions, basketball teams, bomber crews, and tank crews all appear to be more effective if their leader assumes little similarity between the good and bad co-workers. The findings have been replicated several times and there seems little reason to doubt their reliability. But the problem of interpretation is another matter. Is the effective leader one who autistically minimizes similarities between people with whom he can work well and poorly? Or is he accurate in his appraisals of others, while the ineffective leader autistically minimizes differences between good and bad co-workers? A third possibility is that these two kinds of leaders are equally accurate in their evaluations of co-workers but differ from one another in their identifications of good and bad co-workers. In other words, when asked to think of a good co-worker and a bad one, they are guided by different selection criteria. Until we know where the unwarranted assumed similarity lies, we cannot confidently choose between these three explanations which, incidentally, imply somewhat different conceptions of leadership and human interaction.

Some of the difficulties in interpreting assumed similarity scores derive from the fact that this kind of research has generally been of a correlational type. Consequently, there has often been a hen-and-egg problem which could not be resolved by the data. This problem can be illustrated by another study involving assumed similarity between opposites. Uhlmann (8) did a master's thesis which involved obtaining ASo scores for the officers of twenty social fraternities at the University of Illinois. He found that presidents assumed significantly less similarity between a good and a bad co-worker than did vice-presidents, secretaries, or treasurers. On the basis of these data one is

tempted to conclude that the assumed similarity score reflects something about the personality of the individual which is responsible for his being elected president. Indeed, it would not be at all hard to find theoretical and empirical support for this interpretation of the data. But before we conclude that people who assume little similarity between opposites possess unique behavioral propensities which enable them to become fraternity presidents, it is well to take account of another of Uhlmann's findings. About half of the presidents had been vice-presidents, secretaries, or treasurers the year before. The best guess would seem to be that when the current presidents were vice-presidents, secretaries, or treasurers, their assumed similarity scores were about like those of the current incumbents of those offices. In other words, there is a strong suggestion that before these people became fraternity presidents they assumed more similarity between good and bad co-workers than they do now that they are presidents. Perhaps the responsibility and authority which is theirs by virtue of their office has changed their assumed similarity scores. This is a possibility which is being investigated by a new study in which assumed similarity scores are being obtained both before and after election to office.

Of course a considerable portion of the variance in assumed similarity scores may actually be correlated with enduring personality traits or behavioral propensities. Research already completed suggests that such is the case. Steiner and Peters (6) obtained ASo scores for about 50 college students, and several weeks later placed each of these persons in a conformity situation which somewhat resembled the classical experimental situation created by Solomon Asch. However, they used more ambiguous stimuli and only one stooge. Students who had assumed little similarity between a good and a bad co-worker were found to be less likely to conform to the stooge's incorrect judgments. Furthermore, they seemed much more curt and businesslike in their responses to the experimental situation. On the other hand, people who had assumed great similarity between a good and a bad co-worker were not only more likely to conform; they also tended to speak in a soft voice and to hesitate before answering. But after the experiment was over, those who had assumed little similarity between opposites, and who had not conformed, tended to underestimate the number of times they had disagreed with their partner. The data suggest that, having disagreed with their associate, they tended to forget, repress, or deny that disagreements had occurred, thus achieving a sort of interpersonal harmony which other subjects achieved by conforming. It seems apparent that a lot more research must be done before we can have a very secure understanding of the relationships between assumed similarity and human interaction.

SUMMARY

In conclusion, and by way of summary, it seems appropriate to say that in spite of the many negative and contradictory findings which can be cited, research has indicated that interpersonal perception provides a very fruitful approach to the study of human interaction. But many of us have underestimated the difficulties inherent in testing the commonsense propositions which link these two kinds of variables. Only recently have we begun to realize that we shall probably have to struggle at least as hard with this kind of research as we do on any other kind. It is a saddening realization, but a necessary one. Now that we have it, exciting developments should occur.

Manuscript received: October 31, 1958
Revised manuscript received: February 20, 1959

Ivan D. Steiner
Department of Psychology
University of Illinois
Urbana, Illinois

REFERENCES

1. Bruner, J. S., and R. Tagiuri, "The Perception of People," in G. Lindzey, ed., *Handbook of Social Psychology*, Cambridge, Mass.: Addison-Wesley, 1954.
2. Cronbach, L. J., "Processes Affecting Scores on 'Understanding of Others' and 'Assumed Similarity'," *Psychological Bulletin*, 1955, 52, 177–193.
3. Fiedler, F. E., "The Psychological-Distance Dimension in Interpersonal Relations," *Journal of Personality*, 1953, 22, 142–150.
4. Gage, N. L., "Accuracy of Social Perception and Effectiveness in Interpersonal Relationships," *Journal of Personality*, 1953, 22, 128–141.
5. Gage, N. L., and L. J. Cronbach, "Conceptual and Methodological Problems in Interpersonal Perception," *Psychological Review*, 1955, 62, 411–422.
6. Steiner, I. D., and S. C. Peters, "Conformity and the A-B-X Model," *Journal of Personality*, 1958, 26, 229–242.
7. Taft, R., "The Ability to Judge People," *Psychological Bulletin*, 1955, 52, 1–23.
8. Uhlmann, F., "Psychological Distance and Fraternity Leadership," M.A. thesis, University of Illinois, 1958.

Item Reliability and Related Factors in a Community Survey of Emotionality [1, 2]

M. Dean Nefzger, *National Academy of Sciences-National Research Council*
Abraham M. Lilienfeld, *Johns Hopkins School of Hygiene and Public Health*

During 1956 a sample survey of Buffalo and a neighboring suburban town, Kenmore, was carried out to assess the possible relationship of cigarette smoking to emotional status. Integrated into this survey was a study of the reliability of the questions used for determination of emotional status. Because of the increased use of similar questions in epidemiological studies of mental health, we thought there would be sufficient general interest in the results of the reliability phase of this survey to warrant separate reporting (1, 6).

METHODS

General

The sample data were obtained on all adult persons (eighteen years of age and over) in all households of a sample of addresses and on a sample of persons in lodginghouses in Buffalo and Kenmore, New York. Residents of hospitals, convents, dormitories, etc., were not included in the sample. In both the regular addresses and lodginghouses, the samples were selected so that they resulted in a uniform over-all sampling fraction of 1 in 75. A systematic sample of addresses was drawn from the Buffalo City Directory of 1956, and from a list of new building addresses obtained from permits issued by the Bureau of Buildings. In addition, to allow for omissions from the directory, the "half-open interval" method described by Yates (10) was used; addresses thus obtained were added to the original list. A list of lodginghouses was obtained from the Erie County Health Department where each is registered, and a systematic sample was drawn from this list maintaining the same sampling fraction. After selection was completed, sample addresses were arranged into socioeconomic quartiles of the population as determined from census tract data. Addresses were assigned to interviewers so that each had a representation of addresses in all socioecenomic strata.

[1] This investigation was carried out while the authors were on the staff of the Department of Statistics and Epidemiological Research, Roswell Park Memorial Institute, Buffalo, New York, and while the senior author was a Public Health Service Special Research Fellow of the National Institute of Mental Health.

[2] This investigation was aided in part by a research grant CS-9408 from the National Cancer Institute of the National Institutes of Health, Public Health Service.

A total of 4,456 adults was interviewed. About 11 per cent of the adults in the selected sample were not interviewed for various reasons. Some refused, others were too ill to be interviewed, and others could not be contacted despite repeated attempts.

The questionnaire consisted of two forms: The first part, in addition to details of lifetime tobacco use, included identifying data and histories of marriage, occupation, residence, hospitalization, and exposure to medical radiation (diagnostic and therapeutic). Pregnancy and menstrual histories of female respondents were also obtained. This form was completed by the interviewer after oral questioning of the respondent.

The second part, consisting of 43 items, was designed primarily to assess emotional status. However, 12 questions concerned past and present sports participation, religious preference, circumcision among males, and some general aspects of smoking not covered in the first form. Thirty-one questions concerned emotional status; these were taken from a preliminary set used by Star (9) in developing the Neuropsychiatric Screening Adjunct. These 31 items were chosen because the responses of groups of "normals" and "neurotics" to these items were different in the data published by Star.

This second form was read by the respondent who selected and marked his most appropriate answer from among alternatives listed on the form. Twenty-seven of the emotional items had three alternative answers: "often," "sometimes," and "never"; and the remaining four items had four alternative phrases which expressed degrees of frequency in terms appropriate to the content of the question.

Study of Reliability

To investigate reliability, a 10 per cent systematic sample of individuals with completed interviews was selected for reinterview. For practical reasons, the second interview was conducted by a different person; and, unfortunately, the entire 10 per cent sample could not be reinterviewed because of limitations of time. Two hundred and forty-two reinterviews were completed. Consequently, the reinterviews are a 10 per cent sample of approximately half the total sample whose interviews were completed earliest. The average interval between interviews was about 65 days; the frequency distribution of these intervals is presented in Table 1. This report is limited to an analysis of the reliability of answers to the 31 items designed to evaluate emotional status.

RESULTS

As a measure of reliability of response, we were first interested in the percentage of persons who responded identically to individual items on successive

interviews. These percentages are given in Table 2 for items ordered as they were in the questionnaire and are seen to range from 94 (item 3) to 58 (items 9 and 12). A frequency distribution of items by percentage agreement is presented in Table 3 where it is seen that about two thirds of the items had 70 per cent agreement or higher. A crude evaluation of these percentages can be made by comparing them with 33 per cent which represents the agreement to be expected if three alternative responses were selected equally often, and

TABLE 1

Distribution of Intervals Between First and Second Interviews

Intervals in Days	Number of Respondents
11–20	4
21–30	6
31–40	18
41–50	46
51–60	31
61–70	44
71–80	23
81–90	31
91–100	16
101–	23
	242

if responses on successive interviews were independent. The standard error of 33 per cent is 3 per cent for a sample of 242; therefore, all percentages in Table 2 are significantly greater ($P<.01$) than 33 per cent.

To obtain a composite index of emotional status for each respondent, item responses were arbitrarily weighted and summed. Weights of 1 through 3 or 4 (depending on the number of alternatives presented) were assigned to the responses of a single item. The weight of 1 was consistently assigned to the response which intuitively denoted the least anxiety. Weights of either 3 or 4 were assigned to responses denoting most anxiety, and intermediate weights were used for the intermediate alternatives. After summation, the lowest score possible was 31 and the highest score was 97. The higher the score the more anxious or neurotic was the respondent to be regarded. Pearson's r was used to compare the emotion indices from the two interviews. This r between interviews was found to be .73 which is significantly greater than zero but significantly less than the test-retest coefficient reported by Star for the 15 items used in the Neuropsychiatric Screening Adjunct (9).

TABLE 2

Items by Order of Presentation on Questionnaire and Percentages of Identical Responses on Two Interviews

Presentation Order	Item	Per Cent Agreement
1	As far as you know, were you a healthy child or a rather sickly one?	82
2	Did you ever bite your fingernails when you were a child?	83
3*	When you were growing up, did you have any trouble with stuttering or stammering in your speech?	94
4	How would you say people you know feel about you?	68
5	Do you usually like to be by yourself or to be with other people?	83
6	Do you worry very much about things that might happen to you?	61
7	Do you often say things you later wish you had not said?	68
8	How often do people get on your nerves so that you want to do just the opposite of what they want you to do?	59
9	How often does it make you sore to have people tell you what to do?	58
10	Do you ever feel like smashing things for no good reason?	79
11	Do you find that you often have to tell people to mind their own business?	64
12	How often do people hurt your feelings?	58
13	In general, how would you say you feel most of the time, in good spirits or in low spirits?	80
14	Are you ever worried and upset?	67
15	Do you feel that you get more than your share of bad luck?	72
16	Do you ever get so blue and discouraged that you wonder whether anything is worthwhile?	61
17	Do you have any particular physical or health problem?	81
18	Do you often have trouble in getting to sleep or staying asleep?	71
19	Do your hands ever tremble enough to bother you?	83
20	Have you ever had any fainting spells?	87
21	Are you ever bothered by nervousness?	64
22	Have you ever been bothered by your heart beating hard?	78
23	Have you ever been bothered by pressures or pains in the head?	75
24	Have you ever had spells of dizziness?	78
25	Do you bite your fingernails now?	91
26	Have you ever been bothered by shortness of breath when you were not exercising or working hard?	80
27	Are you ever troubled by your hands sweating so that they feel damp and clammy?	76
28	Are you ever troubled with sick headaches?	74
29	How often are you bothered by having an upset stomach?	67
30	Are you ever bothered by having nightmares (dreams that frighten or upset you very much)?	80
31†	Have you ever been troubled by "cold sweats"?	85

* Two items on past and present sports participation preceded this one in the interview; thus, the interview item number of this and every following question was two greater than shown.

† In the schedule ten items followed this one. Eight of these dealt with smoking, one with religious preference, and one with circumcision among males.

Factors Influencing Reliability

In the study of reliability it is of considerable interest to determine whether there are characteristics that distinguish respondents who often agree from those who agree less often in their responses. Knowledge of such differential

TABLE 3

Distribution of Items by Percentage Agreement of Response on Two Interviews

Per Cent Agreement of Response	Number of Items
90–99	2
80–89	10
70–79	8
60–69	8
50–59	3
	31

characteristics would provide a basis for eliminating relatively unreliable groups from final analysis.

To explore this problem, the total sample was divided into three subgroups, depending on the total number of items on which individual members agreed with themselves. One group contained 76 individuals who agreed with themselves on more than 25 items; a second group of 95 individuals agreed on 22 to 25 items; and the remaining 71 individuals agreed with themselves on less than 22 items. The three groups were compared with respect to characteristics obtained on the first part of the interview. Comparisons of qualitative

TABLE 4

Chi-squares, Degrees of Freedom, and Probabilities for Association of Total Item Agreement Groups with Each of Certain Characteristics

Characteristics	Chi-square	df	p*
Social quartile †	5.19	6	.60
Sex	1.54	4	.90
Marital status	5.62	4	.30
Education	18.14	8	.025
Occupation ‡	29.38	18	.05
Industry ‡	17.36	16	.40
Respondent aid §	8.17	4	.10

* The probability assuming no association exists in the population.

† Social quartiles were determined from median contract rentals of census tracts in the 1950 census.

‡ "Occupation" and "Industry" refer to classifications used by the Bureau of the Census in 1950 (11).

§ "Respondent aid" refers to whether the respondent read and marked the interview for himself or was helped in this, and to the combinations of these possibilities for the two administrations.

TABLE 5

Agreement Group Means on Three Measurement Variables

Variable	Number of Agreements			p*
	< 22	22–25	> 25	
Age	41.7	42.7	40.7	> .05
Emotional score †	47.4	43.6	44.4	< .01
Days between interviews	60.0	63.7	70.3	< .025

* Probabilities are those associated with mean-square ratios (F) in independent analysis of variance tests of no difference among means.

† Refers to the first-interview composite of the 31 items.

data were tested by chi-squares which are presented in Table 4, and comparisons of quantitative data were tested by F-ratios which are presented in Table 5.

Table 4 indicates that there were no statistically significant differences between these three groups with respect to social quartile, sex, marital status, industry where employed, and respondent aid. Differences with respect to occupation reached a probability level of .05 and education a level of .025. For these two categories, the detailed distributions of the groups are presented in Tables 6 and 7. Occupational differences, although bordering on statistical significance, do not present a consistent pattern. Educational differences indicate that those who agreed on more items had received, on the whole, more education than the others; yet, the differences are not marked.

TABLE 6

Distribution in Agreement Groups over Broad Occupational Groups

Occupations	Number of Agreements					
	< 22		22–25		> 25	
	Number	Per Cent	Number	Per Cent	Number	Per Cent
Professional, technical, managers, and proprietors	11	15.7	6	6.3	8	10.5
Clerical, sales, craftsmen, and foremen	8	11.4	18	18.9	20	26.3
Operatives, household, service, and laborers	19	27.1	31	32.6	15	19.7
Unemployed	32	45.7	40	42.1	33	43.4
Total	70*	100.0	95	100.0	76	100.0

* Occupation of one respondent unknown.

TABLE 7

Educational Distribution in Three Disagreement Groups

	Number of Agreements					
	< 22		22–25		> 25	
Educational Level	Number	Per Cent	Number	Per Cent	Number	Per Cent
Grade school not completed	19	26.8	18	18.9	11	14.5
Grade school completed	15	21.1	16	16.8	12	15.8
Some high school	16	22.5	33	34.7	19	25.0
High school completed	6	8.4	18	18.9	23	30.3
Some college	15	21.1	10	10.5	11	14.5
Total	71	100.0	95	100.0	76	100.0

With respect to the quantitative variables (Table 5), there are no statistically significant differences in age distribution. The differences in emotional score are statistically significant, although not great; the maximum difference is 3.0. Further, these differences are not consistent; that is, the highest average score of 47.4 was obtained by those who were intermediate in degree of agreement. Such an inconsistent pattern is puzzling and prevents us from drawing any inferences from the comparison. However, it does suggest that respondents with a high "neurotic" score tend to give less reliable answers, which is not unexpected. Obviously, further study of this observation would be of interest. Differences with respect to days between interviews appear to be consistent, that is, greater reliability seems associated with the longer intervals and lower reliability with the shorter intervals. This finding is at complete variance with what one might expect and therefore of considerable interest. Here again, further study should be undertaken.

Another method of analysis of factors that may influence reliability is to compare those whose responses agreed with those whose responses did not agree on individual questionnaire items with respect to the characteristics of interest. This will be recognized as essentially the approach employed by Kendall in the investigation of what she called turnover (5). Here it is applicable only to those items for which the percentage agreement was less than about 70 per cent because the disagreement group becomes quite small for percentages above that value. For qualitative characteristics, such as sex and social quartile, these comparisons were made with chi-square tests; for quantitative characteristics, t tests were used.

Of 77 chi-square tests performed, 6 had probability levels of .05 or less. This result is not too different from what would be expected on a chance basis. Further, no consistent pattern was noted among the small probabilities. Again,

with respect to the quantitative characteristics, there was no striking pattern although several of the t tests were significant at probability levels of .05 or less. Thus, it appears that, with respect to both the qualitative and quantitative characteristics, those who agreed in their responses do not differ from those who did not agree.

DISCUSSION

The proportion of individuals responding identically to individual items is, in general, higher than expected considering the type of question and the procedures involved. There are, however, two major limitations to the interpretation of these proportions. First, it has been argued that such an index of agreement should involve weighting factors proportional to the extent of disagreement (8). However, any simple weights would involve a large element of arbitrariness, and complex weighting procedures could rarely be justified for crude categorical data. Further, in these data, response shifts were predominately of one category so that differential weights would have relatively little influence on the agreement index.

The second limitation concerns the statistical evaluation of even the simple proportions calculated here (3). It was possible to obtain an expected proportion of agreement against which to test observed values, only by making two strong assumptions: equal probabilities for alternate response categories of a given item and independence of successive responses to the same item. The latter is, of course, most relevant to reliability, but the high obtained proportions may indicate only that the former assumption is false. Although it perhaps could be done, the present data are not sufficient to permit isolation of these assumptions.

There are three major alternative explanations for the significantly lower test-retest coefficient obtained in this study as compared with that reported by Star for enlisted men on the Neuropsychiatric Screening Adjunct. One possibility concerns the procedural advantages enjoyed by Star, another relates to questionnaire content, and the third involves scoring procedure. It is, of course, not possible definitely to identify the crucial procedural differences, but among the obvious possibilities are the greater constancy of conditions during group administrations, the greater repeatability of conditions in the Army, the probably greater adverse effect of personality factors in the present study, and the greater variety of experience in our sample during the interval from the first administration to the second. Several other possibilities of this same general class have been discussed by Ellis and Conrad (2).

The second possible reason for the discrepant reliability coefficients derives from the fact that included in our form were some items which were rejected from use in the NSA on the basis of a factor analysis. The rejected items may

have been the least reliable and their inclusion by us may have resulted in a lower test reliability. This possibility is indirectly supported by the observation that among the 16 items rejected by Star, only 6 had agreement percentages of 75 or greater; while among 15 items used by Star, 11 had agreement percentages of 75 or greater. When these frequencies are cast into a contingency table, the chi-square test of independence is found to be significant ($P<.05$). However, the implication concerning the effect on test reliability of using all items does not seem to hold. Scoring response alternatives in the same way as before, a composite score was obtained for the 17 items for which agreement percentages were 75 or greater. The Pearson *r* for the two interview composites thus obtained was only .56, which is almost exactly the value one would expect for the smaller number of items assuming homogeneous item reliabilities (4).

The third alternative explanation of the difference in reliability concerns scoring procedures. After exploring alternative procedures, dichotomous scoring with weights of 0 and 1 was selected for use with the NSA (9). In this study, weights of 1 through 3 or 4 were used. While this is a possible explanation of the reliability differences, it is hardly acceptable in view of Star's findings with simple and differential trichotomous weights and in view of the above evaluation of the effect of item reliability on test reliability.

A specific problem of interest has to do with the *direction* of change among those responses which were not the same on both interviews. A cursory examination was made of only those items with less than 70 per cent agreement, since, when agreement exceeds 70 per cent, the remaining groups are so small as to make analysis of direction of change nearly meaningless.

In general, the great majority of changes in these items were of one step. Relatively few changes were, for example, from a response of "never" to one of "always," and nearly all (more than 90 per cent) were from "never" to "sometimes" or from "sometimes" to "never." In only one item was the proportion of one-step changes less than 90 per cent.

Because of their great preponderance, the assessment of *direction* of shift was restricted to those of one step. With one exception, the items evaluated showed upward (toward a "sometimes" or "always" response on reinterviews) one-step shifts in excess of 50 per cent (the exception is exactly 50 per cent), and in seven of the eleven items the proportions are statistically different from chance. Thus, it appears that on reinterview there is a tendency for those whose responses differ from the intitial interview to answer in what might be considered a more "neurotic" direction. However, these differences are not pronounced.

The analysis of the characteristics of those who agreed with those who did not agree on reinterview indicated that these groups did not differ markedly with respect to age, sex, emotional score, days between interview, occupation,

social quartile, industry, respondent aid, education, and marital status. This result was somewhat surprising since it was expected that such a characteristic as education would show a marked difference. However, this result is consistent with the few reports of similar studies. Palmer studied certain factors that may influence the variability of responses to such questions as marital status, age, employment status, and education (7). She found that variability was higher in districts where there was a relatively high proportion of Negroes or of Negroes and foreign-born white residents. The length of interval between interviews, the quality of interviews (as rated by interviewers), and the amount of interviewer's experience were all unrelated to response variability. In her report, Palmer refers to an unpublished study by Dinkle in which the variability in reported work histories was explored in relation to interviewer personality factors and respondents' ages, education, and nativity. Consistent relations were not found among these variables. Thus, in general, our results are consistent with these other studies. Needless to say, the lack of such relationships is comforting from a methodological viewpoint.

SUMMARY

In the course of a sample survey of 4,456 adults in the Buffalo, New York, area, a reliability study was made of the questionnaire used to assess emotional status. Two hundred forty-two respondents were reinterviewed after various intervals of time. These 242 respondents represent a 10 per cent systematic sample of approximately half the total sample whose original interviews were completed earliest.

An evaluation was first made of the percentage of persons responding identically on both interviews to each given item. The agreements range from 94 to 58 per cent, and all are significantly greater than a theoretically expected 33 per cent.

By applying simple, arbitrary numerical weights to item alternatives, composite emotional indices were obtained for all respondents. The Pearson r between the two sets of scores is .73, which is significantly greater than zero but significantly less than a test-retest coefficient previously reported for a similar set of items.

An effort was also made to isolate factors which characterize groups of individuals who agree more or less often. Such factors as age, sex, socioeconomic quartile, marital status, and industry of employment were found unrelated to extent of agreement. Occupational, educational, and emotional score differences are statistically significant, but are either small or difficult to interpret. Perhaps the most interesting finding is the significant inverse relation of days between interviews to degree of unreliability.

Groups agreeing and disagreeing on individual items were also compared

on a variety of factors. Several significant statistics were obtained, but because of the large number computed, the over-all result is not in excess of expectation.

Alternative procedures for computing agreement indices and difficulties encountered in evaluating even the simplest index are briefly discussed. Where differences are known to exist, an attempt is made to reconcile our results with those of earlier studies. Reports of results similar to ours are also pointed out. In discussing response changes, it is noted that most were of only one step or category and were predominantly in an upward direction (i.e., toward the more "neurotic" end of the scale).

Manuscript received: November 20, 1958
Revised manuscript received: March 2, 1959

M. Dean Nefzger
Division of Medical Sciences
National Academy of Sciences–National Research Council
2101 Constitution Avenue
Washington 25, D. C.

REFERENCES

1. Artesen, K., "An Investigation of the Questionnaire Method by Means of the Cornell Index (Form N2)," *Acta Psychiatrica et Neurologica Scandinavica,* 1957, 32, 231–279.
2. Ellis, A., and H. S. Conrad, "The Validity of Personality Inventories in Military Practice," *Psychological Bulletin,* 1948, 45, 385–426.
3. Goodman, L. A., and W. H. Kruskal, "Measures of Association for Cross Classification," *Journal of the American Statistical Association,* 1954, 49, 732–764.
4. Gulliksen, H., *Theory of Mental Tests,* New York: Wiley, 1950.
5. Kendall, P., *Conflict and Mood: Factors Affecting Stability of Response,* Glencoe, Ill.: Free Press, 1954.
6. Macmillan, A. M., "The Health Opinion Survey: Technique for Estimating Prevalence of Psychoneurotic and Related Types of Disorder in Communities," *Psychological Reports,* 1957, 3, 325–339.
7. Palmer, G. L., "Factors in the Variability of Response in Enumeration Studies," *Journal of the American Statistical Association,* 1943, 38, 143–152.
8. Robinson, W. S., "The Statistical Measurement of Agreement," *American Sociological Review,* 1957, 22, 17–25.
9. Star, S. A., "The Screening of Psychoneurotics in the Army," in S. A. Stouffer *et al., Studies in Social Psychology in World War II: Measurement and Prediction* (Vol. 4), Princeton, N. J.: Princeton University Press, 1950.
10. Yates, F., *Sampling Methods for Censuses and Surveys,* 2d ed., New York: Hafner, 1953.
11. U. S. Department of Commerce, Bureau of the Census, *1950 Census of Population: Alphabetical Index of Occupations and Industries,* Washington, D. C.: U. S. Government Printing Office, 1950.

Studies in Personal Space[1]

ROBERT SOMMER, *The Saskatchewan Hospital*

Surprisingly little is known about the way people use space. Social scientists in the field of human ecology have been concerned primarily with the distribution of such things as social classes, economic institutions, and mental illness. An almost unexplored area is microecology or the way that people in pairs or small groups arrange themselves.

In the studies of Hediger (5), Howard (6), and Von Uexkull (15), "space" has had two different meanings. The more familiar of these refers to space in the geographic sense, i.e., space as area. It is most commonly discussed with reference to the animal's territory or home. However, some writers have applied the concept of territory to human behavior. W. F. Whyte (16) and Thrasher (14) have studied the territories of adolescent gangs. W. H. Whyte (17) also has studied the groupings of people within a housing development. Probably the most intensive investigation thus far has been that of Barker and Wright (2) in the Midwest. Yet this still remains relatively unexplored territory for social scientists.

The second way in which the term "space" is used can be called "personal space of the organism." Although it has its roots in the work of zoologists and ethologists, it is an entirely different concept from that of territory. Personal distance is the distance that the organism customarily places between itself and other organisms. This distance may vary from species to species and individual to individual. Hediger speaks of this as "flight distance" and has measured this for hundreds of animals.

This concept would seem to have relevance for the study of human behavior, although it has never been studied empirically so far as we know. It seems obvious that people feel uncomfortable when they talk to others who either stand too close or too far away. The concept itself has been used several times in the literature. David Katz (8) used the term "personal space" and compared it to the shell of a snail. Von Uexkull used the graphic analogy of people "surrounded by soap-bubble worlds." Stern (13) developed the concept of personal world. He noted that the physical world was without a center, but the

[1] This study was supported by grants from the Rockefeller Foundation, the Province of Saskatchewan, and the Department of Health and Welfare (Ottawa). Drs. Humphry Osmond and T. E. Weckowicz gave valuable advice and encouragement. The following people were research assistants during the study: R. Dewar, R. Hall, Eleanor Fenton, Sheila Kelly, Gwen Witney, Dorothy Sommer, Sandra Wright.

personal world had a natural center from which and toward which everything pertaining to it extends. This center is the person himself about whom the personal world is oriented. Stern's analogy to the "snail shell" of Katz and the "soap bubble" of Von Uexkull was to describe the "personally near" as an "aura" surrounding the person.

The concepts of "personal space" can be distinguished from that of "territory" in several ways. The most important difference is that personal space is carried around while territory is relatively stationary. The animal or man will usually mark the boundaries of his territory so that they are visible to others, but the boundaries of personal space are invisible. Personal space has the body as its center, while territory does not. Often the center of territory is the home of the animal or man. Animals will usually fight to maintain dominion over their territory but will withdraw if others intrude into their personal space.

This paper will describe several studies of personal distance. The first study was purely observational. That is, we were interested in the way that people who were *already interacting* were arranged. Subsequent studies used experimental procedures in that Ss were asked to interact, and observations were made as to how they arranged themselves. Schizophrenic patients are also used as subjects. We hope to learn if their difficulties in communicating with others are reflected in the way they place themselves. Will they sit too close to the other person or too far away? We are also interested in possible sex differences in seating arrangements. In all studies, every effort was made to secure natural conditions, i.e., to locate the experiments in settings where interaction was customary and where the subjects would feel comfortable.

Experiment One

Method

The first study was conducted in the staff dining hall of a 1500-bed mental hospital. Most staff members eat in this hall, hence it serves nurses, plumbers, accountants, secretaries, etc. The hall was 36 by 68 feet and contained thirteen tables (36 by 72 inches). Each of these tables consisted of two 36 by 36-inch tables placed together and with eight chairs arranged around it. One of these tables, along with the chairs (lettered for identification) is shown in Figure 1. Service in the dining hall was cafeteria style. The staff members secured their plates at the front of the cafeteria, went in line, were served, and then sat at whichever table they chose.

In this study we were interested in the chairs occupied by people who were interacting. That is, did communication occur between the head and foot positions (A-E in Figure 1), or between people seated alongside of one another

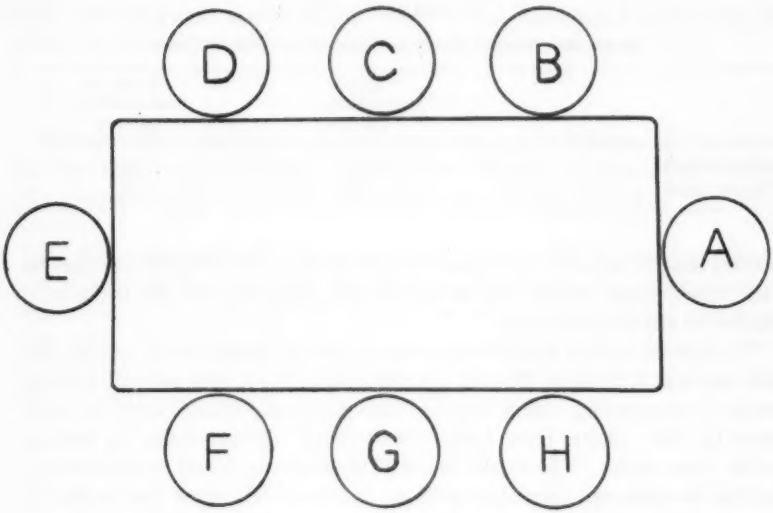


FIGURE 1

Table and Arrangement of Chairs Used in Experiment

(B-C or G-H)? This study was purely observational and, in one sense, exploratory in that no specific hypotheses were formulated.

Two observers were used throughout the study and all observations took place during the noon meal. Both O's would secure their food and then select a table at which to sit. On some days the O's sat at the front of the hall, sometimes at the rear, etc. The O's would then look around, and if both agreed that they had a clear view of another table, they would wait two minutes, and then record five seconds of interactions.

The recording was done on mimeographed forms, similar to Figure 1. Each would independently record the verbal interactions that ensued during the five-second period. If no interactions occurred, another table was selected later. Usually two records were made (by each O) at a meal, and the observations were done over a two-month period. In all, a total of 50 observations was made by each O.

RESULTS

The 50 interaction records for each observer were summed to find the number of conversations between people sitting in all possible combinations of chairs (A-B, A-C, F-G, F-H, etc.). The totals for Observer One were correlated with the totals for the second observer. The Pearson r between the two ob-

TABLE 1

Interactions between People in Adjacent and Distant Chairs

	Chance expectancy, per cent	Results of 50 observations, per cent
Adjacent chairs	39	73
Distant chairs	61	27

servers was .91, which is significant beyond the .01 level. This showed that the observations were reliable. In the rest of this paper we will use the average figures for the two observers.

The first distinction that can be made is between chairs next to one another and chairs at a distance. That is, communication is not only possible between people in neighboring chairs, but also between people whose chairs are separated by other chairs. These latter can be called "distant chairs" in contrast to the "near chairs." The results of the 50 observations, based on the averages for the two O's, are presented in Table 1. These data show that of the 67 recorded interactions, there were only 18 interactions between people sitting in distant chairs. This differs markedly from the figures based on the permutations and shows that the people interacted with those sitting in neighboring chairs. This is hardly unexpected, but it is necessary to recognize this trend first before making a separate analysis of the interactions between people in neighboring chairs.

If all permutations of possible interactions between *neighboring* chairs are computed, we should predict, on the basis of chance, 36 per cent of the interactions between people sitting *side by side*, 27 per cent between people sitting *face to face*, and 36 per cent between people sitting *corner to corner*. Both the expected frequencies and the obtained frequencies from the 50 observations are presented in Table 2.

The results in Table 2 show that the interactions between people seated corner to corner (i.e., A-B, A-H, E-D, E-F) exceeded chance expectancy while

TABLE 2

Arrangements of People Interacting

Arrangement	No. of pairs in this arrangement	Chance expectancy
Corner-to-corner	30	18.5
Side-by-side	16	18.5
Face-to-face	5	13.9

those between people seated face to face were less than chance expectancy. A test of significance of these data yields a chi-square of 19.1, $p < .01$.

DISCUSSION

These results parallel those of Festinger, *et al.* (3) and James (7) in showing that communication tends to take place between neighbors. Our results also showed that the corners of the table were the loci of most of the interactions.

We felt that these results made it necessary to determine if corner-to-corner interaction would predominate in other size tables, and also to see whether people who desire to interact will choose to sit corner to corner when presented with other possibilities. That is, there is the possibility that our results were not so much due to the corner positions sparking the interaction as to people who wanted to talk seating themselves in the corner positions.

Experiment Two

In this study we hoped to learn the way people arrange themselves when they desire to interact. Our method correspondingly changed from natural observation to active experimentation. Our goal was to ask groups of subjects to interact and then observe the ways in which they arranged themselves.

PROCEDURE

In order to conceal the purpose of the experiment from the subjects, we asked them to discuss particular proverbs and told them that they would be interviewed as to the meaning of the proverbs. The subjects were ushered to the door of the cafeteria and told:

> This is a study of discussion groups and the way people discuss things. On the card here is a statement I would like you to discuss. Please go into the cafeteria, sit down, and begin discussing it. After a while you will be interviewed as to what you discussed.

Various groups were used under different conditions. These will be described in detail in the next section. All subjects were interviewed as to the meaning of the proverb. During the interview E recorded how the people were sitting. It was apparent that none of the subjects suspected the purpose of the experiment.

RESULTS

A. In this part of the study six pairs of subjects were asked to sit at the type of table shown in Figure 1 and discuss the proverbs. The subjects were assorted employees (secretaries, personnel workers, nurses, etc.) and the groups were homogeneous as to sex. All six pairs arranged themselves at the corners (A-B,

A-H, E-D, or E-F). On the basis of permutations of all possible arrangements, this result is highly significant ($p < .01$).[2]

B. The same procedure was repeated in another cafeteria which contained tables 35 by 48 inches. There were four chairs around each table. The subjects used were ten pairs of hospital employees, again a heterogenous group of hospital employees (accountants, nurses, recreation workers, etc.). The results showed that eight of the pairs arranged themselves corner to corner, while two arranged themselves face to face.

C. The experiment was repeated in the staff dining hall (used in Study One and Study Two, Part A), but only single tables (36 by 36 inches) were used. The subjects could either sit face to face or corner to corner. The subjects used were nine pairs of student nurses from a nearby general hospital. The results showed that eight of the pairs arranged themselves corner to corner while one pair chose a face-to-face arrangement.

D. The experiment was repeated again, with groups of three subjects each. This time the topic discussed was "If you could help design a mental hospital ward, how many people would you prefer in a bedroom?" The subjects were eleven groups of three persons each, all members of a mental health association of a distant city. The results showed that nine of the groups chose the corner positions (H-A-B or D-E-F) while two groups chose other arrangements (H-B-D and C-D-E).

E. In view of the many discussions of the difficulties that schizophrenic patients have in communicating with others, the procedure was repeated using some of our hospital patients. We hoped to learn if the schizophrenic patients would avail themselves of the corner-to-corner arrangement chosen by the other subjects.

The subjects in this study were 26 pairs of schizophrenic patients in this hospital. Most had been used in other researches and had been reasonably cooperative. Nine of the S's were used twice, but the second time each was paired with a new partner. There were 38 males and 5 females in the sample and the average age of the group was 38.9 years.

As a control group, 11 pairs of nonschizophrenic mental patients in this hospital (depressives, alcoholics, psychopaths, etc.) were tested.

The subjects were taken in like-sex pairs to the entrance of the cafeteria, given cards containing a proverb, and asked to go inside, sit down, and discuss it.

The results are presented in Table 3. It shows that the nonschizophrenic patients resembled the normal group in that they used the corner positions

[2] The significances are evaluated by two-tailed chi-square tests, except where the frequencies are too small. Then Fisher's exact methods are employed.

more than would be expected by chance ($p < .05$). The schizophrenics, however, did not show any such trend. The greatest number chose to sit in "distant" arrangements. None of the normal or nonschizophrenic groups used this type of arrangement.

Observation of the session disclosed that very few of the schizophrenics talked at all. Most had wandered aimlessly up the aisles and sat down anywhere. They would face the floor or away from their partner. They seemed to lack both direction and interest.

TABLE 3

Seating Patterns of Pairs of Patients

	Opposite	Corner	Side	Distant
Schizophrenics	8	4	4	10
Nonschizophrenics	3	7	1	0

Experiment Three

Method

The studies that follow all employed a "decoy." This was a person who was a confederate of E and who was *already seated* at a particular chair *before* the subject entered the room. The subject would then be asked to go inside and discuss a topic with the other person (i.e., with the seated decoy). The decoy was instructed to sit passively and look down at the typewritten card until S began to sit down. Then the decoy would look up, greet S, and begin discussing the topic.

The normal Ss used in the study were a heterogeneous group of hospital employees and visitors. Many of the sample were used in two or more procedures, e.g., both with a male decoy and with a female decoy. No S discerned the actual purpose of the study. The seating usually took place automatically, i.e., without conscious decision. A few of the patients had to be urged to sit down, but this was very rare in the normal group. Most of the decoys were members of the Research Department, although several other staff members were used. In almost all conditions, at least two decoys of each sex were used.

The study is divided into three parts, depending upon the seat occupied by the decoy.

Part 1. Decoy at *B*[3]

Part 2. Decoy at *C*

Part 3. Decoy at *B*—no corner chair (i.e., Chair A removed)

[3] The chairs will be identified according to the letters shown in Fig. 1.

The results are presented in terms of the relationship between the chair selected by S and the chair occupied by the decoy. If the decoy was at *B*, the subject could sit alongside him (*C*), facing him (*H*), at the corner (*A*), or at one of the "distant" chairs (*D, E, F, G*). The schizophrenic Ss were all patients at this hospital.

RESULTS

Part 1: Decoy at B

The results for this condition are summarized in Table 4. The data are classified as to (a) the sex of the decoy; (b) the sex of the S; and (c) whether the Ss were normal or schizophrenic.

It can be seen that there are some striking differences in the positions occupied by the Ss. With the female decoy the female (normal) Ss tended to sit in the corner position. Males were more inclined to sit opposite both male and female decoys. The results from the female Ss with the female decoy differ significantly at beyond the .05 level from the other three groups.

There is a marked difference ($p < .01$) between the normals and the schizophrenics in this condition. The schizophrenics sat alongside the decoy far more than the normals and made almost no use of the corner chair (*A*).

TABLE 4

Seats Chosen When Decoy Sits at Seat B

Group	Sex of Ss	Sex of decoy	Number of Ss choosing to sit: Alongside (C,D)	Opposite (H)	Corner (A)	Opposite side (F,G)	Distant corner (E)	N
Normals	Male	Male	1	10	5	0	0	16
		Female	0	11	5	1	1	18
	Female	Male	0	11	6	0	0	17
		Female	2	0	10	0	0	12
Schizophrenics	Male	Male	3	6	1	1	1	12
		Female	4	8	1	1	0	14
	Female	Male	6	5	0	1	0	12
		Female	12	0	1	0	1	14

Part 2: Decoy at C

The results when the decoy sat at *C* are summarized in Table 5. Again the females sit "closer" to the female decoy, although this time their preference is heavily weighted toward sitting alongside her. Only one of the normal males sat alongside a decoy. This difference is significant beyond the .01 level. The

TABLE 5

Seats Chosen When Decoy Sits at Seat C

Group	Sex of Ss	Sex of decoy	Number of Ss choosing to sit: Alongside (B,D)	Opposite (G)	Corner (A,E)	Opposite side (F,H)	N
Normals	Male	Male	0	4	4	3	11
		Female	1	10	8	4	23
	Female	Male	1	10	4	2	17
		Female	13	1	4	4	22
Schizophrenics	Male	Male	7	2	6	3	18
		Female	2	4	5	2	13
	Female	Male	3	6	1	1	11
		Female	8	3	5	1	17

only significant difference between the normals and the schizophrenics in this table is that seven of the male schizophrenics sat alongside the male decoy while none of the male normals did.

Part 3: Decoy at B with No Corner Chair

The inclusion of this condition was dictated by our curiosity as to whether face-to-face or side-by-side seating would predominate if corner seating were precluded. Thus the neighboring corner chair (*A*) was removed but otherwise the procedure was identical to that of previous sessions. As only three of the 52 schizophrenics had occupied the corner chair when it was present (in Part 1), there seemed no point in assessing their reaction if it were absent. Hence only normal Ss were used in this condition.

The results in Table 6 disclose the same trend for females to prefer sitting alongside the female decoy significantly more than males, or than either sex with a male decoy.

Part 4: Some Further Studies

One surprising aspect of these results was that no trace was found of the "distant" seating pattern prevalent among the schizophrenic Ss of the previous study. "Distant" seating refers to one person sitting in the chair that is separated from his partner by another chair. In the previous study 10 of the 26 pairs of schizophrenics had arranged themselves in this way. As many of

TABLE 6

Seats Chosen When Decoy Sits at Seat B (No Corner Chair)

Sex of Ss	Sex of decoy	Number of Ss choosing to sit: Alongside (C,D)	Opposite (H)	Opposite side (F.G)	Distant corner (E)	N
Male	Male	1	10	0	0	11
	Female	1	9	0	0	10
Female	Male	3	14	0	0	17
	Female	8	8	0	0	16

these Ss were also used in the present study, we had expected to find similar results. Yet the data showed the Ss tended to sit *too near* rather than too far from the decoy.

However, the present procedure was very different from that of the previous study. Previously pairs of Ss had been escorted to the door of the cafeteria, given the topic, and asked to go inside, sit down, and discuss it. In this study, *single* Ss were brought to the door of the cafeteria, given the topic, and asked to go inside, sit down, and discuss it with a decoy who was already seated. The major differences were that the decoy was a staff member whereas the partner had been a patient, and the S was now given a clear "target" (the seated decoy) toward which he could orient himself.

The procedure of the present study was thus repeated but with schizophrenic decoys. We were interested in learning whether the patients would sit differently vis-à-vis a schizophrenic patient than they had vis-à-vis a staff member. Two separate experiments were carried out, one with a male decoy at seat *B* and another with a male decoy at seat *C*. To simplify matters only male decoys and male Ss were used. When the decoy sat at *B*, 11 Ss and 3 separate decoys were used. When the decoy sat at *C*, 18 Ss and 4 decoys were used.

The results are summarized in Table 7 and show that the status of the decoy as patient or nonpatient had very little influence on S's choice of a seat. One third of the patients sat side by side (which was very rare among the normal males) and no trend toward "distant" seating was found.

However, unlike the previous procedure these Ss had now possessed a clear "target" or goal (the seated decoy) toward which they could orient themselves. Previously neither S had known where his partner was going to sit nor what was expected of him. This may account for the aimless wandering and the "distant" seating of Study Two. Other writers (4) have commented on the

TABLE 7

Seating Patterns of Schizophrenic Ss with Schizophrenic Decoys

	S chose to sit:					
	Alongside	Opposite	Corner	Opposite side	Distant corner	N
Decoy at B	6	3	1	0	1	11
Decoy at C	4	6	5	3	—	18
Total	10	9	6	3	1	29

importance of a clearly structured situation for schizophrenic patients. If they do not know what is expected of them, their performance deteriorates and any group structure dissolves. This seems to have been what occurred when the two schizophrenics were asked to discuss a topic together. When a decoy was present, the S could be reasonably confident that if anything were expected of him, the decoy would let him know. No such guidepost existed when two schizophrenics were together. As research on mental hospital wards has shown (10, 11), spontaneous conversations between chronic schizophrenic patients are very infrequent occurrences. Many of these patients will, however, talk to nurses if they are approached.

DISCUSSION

The observations of interactions in Part One show that in small groups there is no simple relationship between distance and communication. Subjects who were sitting side by side were physically closer to one another but interacted less than subjects sitting corner to corner. However, the trend in all the data is that people sitting in neighboring chairs (regardless of their positions) will be more likely to interact than people sitting in distant chairs. Principles governing spatial arrangements in small groups must, therefore, take into account both the distance between people and their positions vis-à-vis each other. Steinzor's (12) hypothesis that interaction will be more likely between

people who can see one another is supported by these data but must be qualified in terms of the angle between the participants. The subjects in these studies did not show any preference for face-to-face seating, where corner seating was possible.

There are several conclusions to be drawn from the sessions in which the decoy was used. Females will sit closer to a female than to a male. This is closer than males will sit to decoys of either sex. Of the normal Ss only the females chose to sit alongside the decoy. The males overwhelmingly preferred the chair opposite the decoy. This result parallels the observation that females in our culture will often be seen holding hands or kissing other females, whereas these behaviors are uncommon for males.

Obviously, there are cultural influences at work here. Perhaps if Fiji Islanders were used as subjects they would prefer sitting on the tables instead of on the chairs. Yet this does not diminish the relevance of this research for our own culture. The situation has parallels in animal research where rats raised in one laboratory may be more gregarious or aggressive than rats raised in another laboratory. Domesticated dogs or horses are undoubtedly very unlike their wild brethren, especially in regards to their "flight distance" (cf. Hediger), but this does not deny the importance of studying the habits of domestic animals. In fact, there are more practical reasons for studying the habits and needs of captive and domestic animals than there are for studying the habits and needs of wild animals. The same holds true in the study of human ecology. As long as man must live in a world of walls, furniture, doors, and fences, there is good reason to study how they influence his behavior.

The second conclusion from these data is that schizophrenic patients have an impaired concept of personal distance. This was especially evident when the schizophrenics of both sexes sat alongside the male decoy. This rarely happened in the normal group. One can only speculate how this affects the relations between the schizophrenic and his neighbors. He undoubtedly intrudes on the personal space of those around him. This may cause the normal person, whether a nurse or a relative, to draw away from him (i.e., to maintain the customary personal distance). This may leave the schizophrenic feeling rejected by the withdrawal and the nurse feeling uncomfortable at the intrusion into her personal space.

It is possible to view this impairment of personal distance as a social artifact, but this time as a product of the mental hospital society. Usually a mental patient has little or no privacy and very few things he can call his own. The dayroom, dining hall, and especially the dormitory are often overcrowded. This may produce a situation analogous to that of the canaries in the small cage reported by Allee (1). When the area was crowded, no territories were staked out, but upon transfer to a larger cage, the birds began to

select areas of their own. It would be interesting to observe in such a situation whether prolonged imprisonment in a small cage would destroy or impair the birds' desire for individual territories.

It is curious that some of the most valuable research on the way that people or animals use space has occurred in situations of maximum constraint. Hediger's work with animals in zoos and circuses is exemplary. There is also a growing literature on the ecology of mental hospital wards and geriatrics centers (9, 10, 11). Studies of human interaction in fields or courtyards are practically nonexistent.

In closing it should be mentioned that our reason for conducting the studies in cafeterias was not simply that other space was unavailable. Rather we hoped to achieve a setting where interaction would be relaxed and natural. We wanted to avoid, as much as possible, the use of a cold and impersonal laboratory, where our subjects might suspect hidden microphones and one-way mirrors. Following the methods of the animal biologists, we hoped for a setting where free interaction could occur. Our feeling is that the cafeterias were successful in getting the subjects to feel at ease. For example, the director of a group of nurses that were subjects commented spontaneously, "The girls are at home discussing things in a cafeteria."

SUMMARY

This was a series of studies of the ecology of small discussion groups. The central question was the way that people will arrange themselves when they are interacting. The setting of the study was a cafeteria in a large mental hospital, selected because it was a place where interaction could be free and "natural." The chairs around the rectangular tables were used as coordinates for locating the positions of the Ss. Ss could be described as sitting opposite one another, alongside one another, at corner positions, or at some distance from each other. The study consisted of several related investigations.

1. Observations were made during the noon meal as to the position at the table of the people who were talking. The results disclosed that people in neighboring chairs interacted more than people in distant chairs. Also those in corner positions interacted more than people alongside one another or facing each other.

2. Pairs of Ss were asked to enter the cafeteria and discuss various topics. Again a preference for the corner positions was seen. The same trend was evident when groups of three Ss were used. Schizophrenic Ss made considerable use of "distant" arrangements and would face away from their partners and refuse to speak. However, nonschizophrenic mental patients resemble the normal group in their preference for corner positions.

3. In this study, a "decoy" was used. This was a person who was already seated at a table before S entered the room. The S was then asked to walk over and sit down and discuss a topic with him. Decoys and Ss of both sexes were used in various combinations (i.e., male Ss with female decoys, male Ss with male decoys, etc.). The results disclosed that females would sit "closer" to female decoys than they would to male decoys; this was also "closer" than males would sit to decoys of either sex.

Schizophrenics Ss were also tested, and they appeared to have an impaired concept of social distance. That is, male patients would sit alongside a male decoy, female patients would sit alongside a male decoy, etc. This hardly ever occurred in the normal group.

Manuscript received: October 27, 1958
Revised manuscript received: March 30, 1959

Robert Sommer
Saskatchewan Hospital
Weyburn, Saskatchewan
Canada

REFERENCES

1. Allee, W. C., *The Social Life of Animals*. New York: Norton, 1938.
2. Barker, R. G., and H. F. Wright, *Midwest and Its Children*. Evanston, Ill.: Row, Peterson, 1957.
3. Festinger, L., S. Schachter, and K. Back, *Social Pressures in Informal Groups*. New York: Harper, 1950.
4. Goodrich, D. W., L. Mazer, and B. Cline, "Fostering the Involvement of the Psychiatric Patient in Group Activities, *Psychiatry*, 1958, 21, 259–268.
5. Hediger, H., *Studies of the Psychology and Behavior of Animals in Zoos and Circuses*. New York: Criterion, 1955.
6. Howard, E., *Territory in Bird Life*. London: John Murray, 1920.
7. James, J., "Some Elements in a Theory of Small Groups," *Research Studies of the State College of Washington*, 1950, 18, 150–151.
8. Katz, David, *Animals and Men*. New York: Longmans, Green. 1937.
9. Pace, R. E., "Situational Therapy," *Journal of Personality*, 1957, 25, 578–588.
10. Sommer, R., "Social Interaction on a Geriatrics Ward," *International Journal of Social Psychiatry*, 1958, 4, 128–133.
11. Spohn, H. E., *et al.*, "Social Psychological Research Approaches in the Study of Schizophrenia." Paper presented at APA meeting, Washington, D. C., 1958.
12. Steinzor, B., "The Spatial Factor in Face-to-Face Discussion Groups," in *Small Groups*, ed. by P. Hare, E. Borgatta, and R. F. Bales. New York: Knopf, 1955.
13. Stern, W., *General Psychology*, translated by H. D. Spoerl. New York: Macmillan, 1938.
14. Thrasher, F. M., *The Gang*. Chicago: University of Chicago Press, 1927.
15. Von Uexkull, J., "A Stroll through the Worlds of Animals and Men, in *Instinctive Behavior*, Claire Schiller, ed. New York: International Universities Press, 1957.
16. Whyte, W. F., *Street Corner Society*. Chicago: University of Chicago Press, 1943.
17. Whyte, W. H., *The Organization Man*. New York: Simon and Schuster, 1956.

Effects of Being Witty on Position in the Social Structure of a Small Group [1]

JACQUELINE D. GOODCHILDS,[2] *Cornell University*

One outstanding characteristic of interaction in small groups is the tendency for humorously toned conversation to bubble up again and again throughout the proceedings. Aside from essentially casual speculation, however, this phenomenon has not been directly examined either theoretically or experimentally. Studies focused on interaction process and/or content have either ignored the fact that communications may be couched in humorous terms or have a priori assigned some constant meaning, e.g., tension release, to such occurrences (see Bales [1; 5, p. 123] for a conscientious effort to deal with this problem). Studies directly concerned with humor per se have tended to consider social aspects, if at all, only in connection with individual differences in responsiveness to jokes (2). The present study is an attempt to deal directly with one aspect of this problem: the initiation of humor (being witty) in a social setting and the effects of such behavior on the position of the Wit in the social structure of his group. Witty individuals, restricted entirely to verbal expressions of humor, were selected according to their characteristic kind of humor to represent polar types: Sarcastic Wits whose jokes are predominantly of the biting, ridiculing, pointed variety and Clowning Wits whose humor is primarily whimsical, silly, or frivolous. Social position was examined on two dimensions of structure: power to influence and popularity. The hypothesized relationships were: a Sarcastic Wit has power to influence others but is an unpopular person; a Clowning Wit, on the other hand, is popular but has little power.

Of the numerous factors which might be expected to disturb the posited relationships some were eliminated by the already mentioned restrictions on type and consistency of humor and by the use of a standard nontask-oriented social setting, a situation that is clearly not inappropriate for humor. Additionally, however, in order to provide the experimental variable "being witty" with the relatively clear field required for an initial investigation, roles or qualities known to be important determinants of social status had in some way to be brought under control. These considerations dictated the use of artificial simulated interaction in the form of written records of conversations of fic-

[1] Based upon a doctoral dissertation submitted to the Department of Psychology, Cornell University.

[2] Now with Fels Group Dynamics Center, University of Delaware.

tional small groups. If it is demonstrated that written protocols meeting the several requirements were perceived by subjects similar to the protocol group members as in fact revealing individuals being witty in the various ways, then the original hypotheses can be tested by having additional subjects read the protocols and describe their perceptions of the group social structure along the two dimensions of concern.

The testing situation may be conceived of as two steps removed from the original, which of course would have involved obtaining measures of the effects of being witty from participating members of the interacting group. A one-step removal would have consisted of measures from observers of an actual interacting group; the further remove here was that the responding subjects had only a written record of the interaction on which to base their judgments. To the extent that observation differs from participation and that "observation" of a written script differs from observation of an actual interaction situation the hypothesis testing is approximate.

STIMULUS PROTOCOLS

Twenty-four fictional written protocols were developed, each involving a casual *entr'acte* conversation by three individuals of the same sex, described as college student classmates. Explicit references to internal or external differentials of prestige, abilities, popularity, power, or access to relevant resources were omitted. The only systematic difference among group members within a protocol was humor: one of the three actors consistently attempted to be amusing. With the exceptions that this member (the Wit) never began and always closed a protocol, his interactions were distributed in no systematic way throughout the conversation. Within each protocol the who-to-whom pattern was balanced: i.e., each member preceded and followed each other member an equal number of times, and contributions were generally directed to the group as a whole. One half of the protocols involved 6 contributions per actor; the others involved 12; the number of words contributed by each of the three actors was roughly one third of the total. Of the 24 protocols, 12 contained a Wit whose humor tended to be Clowning; the other 12 Wits were relatively Sarcastic. There were 6 'short' and 6 'long' protocols of each humor type. The protocols were designed so that the actors might be presented either as male or female, and each actor in a protocol was assigned two alternate names, one male and one female, at random from two lists of first names taken from Webster.

In a regular classroom situation each of 49 men and 24 women [3] received a

[3] The total 223 subjects were Cornell undergraduates, 65% in their first two years of study, over 90% between the ages of eighteen and twenty-five.

booklet containing the 24 protocols (with the actors represented in each case as of the same sex as the responding subject) arranged in one constant order—random, subject to the restriction that in each group of 6 there be 3 of each humor type. Subjects were instructed to read through the protocols, circling those remarks "which you would consider amusing, clever, witty, funny, humorous if overheard in an impromptu casual conversation." After completing this first task, the subject identified (by writing in the name on an answer sheet) that person in each protocol who "consistently attempts to be humorous" and finally classified the humor by type. Descriptions of the two types of humor—condensed versions of the category descriptions of Bales (1, p. 179 and p. 194)—were presented and the subject selected from his list of 24 attempted Wits the "twelve persons whose attempts to be humorous best fit the A [Sarcastic] description." [4]

Responses to the humor questions yielded four scores for each protocol: (a) Wit identification—the percentage agreement on the identity of the one person attempting to be amusing; (b) humor typing—the percentage agreement on the classification of the Wit's humor as Clowning or Sarcastic; (c) amount of humor—average number of remarks circled as humorous; and (d) Wit differentiation—an average of the number of circled remarks by the Wit minus the number of circled remarks by non-Wits. Calculations were done separately for each sex, although comparisons across sex on these measures indicated that men and women reacted similarly to the humor in the stimulus protocols (statistically stable positive correlations between the sexes on all four indices). A different experimental treatment (Variation I or II) was associated with each subject sex.

Variation I: Males

Within the type and length categories, 12 protocols were selected from the original 24 so as to maximize the four criterion scores and maintain equivalence of score for the two humor types. For the chosen protocols, male subjects achieved 94 per cent agreement on Wit identification and 78 per cent agreement on humor typing. The amount of humor averaged 2.27 remarks per protocol; the Wit differentiation score was 1.80 (roughly one laugh per minute in terms of time to act out a conversation). There were no statistically significant differences [5] on the four indices between the six Sarcastic and the six Clowning protocols.

[4] The labels 'Clowning' and 'Sarcastic' were never used with subjects; they are used in this report only for convenience in discussing the typology.

[5] 'Significant' throughout this report shall be taken to mean at or beyond the 5% level, two-tailed test.

Variation II: Females

The interest in this treatment was—unlike Variation I where the clarity of the humor was maximized by selection from the total—in retaining all protocols with the full range of degree of consistency of humor type and of degree of success in attempted humor by the Wit.

A random division of the 24 protocols into two groups of 12 was made, subject to the restriction that each group contain 6 of each humor type. Within each group, the protocols were combined by type into the four most Clowning, the four most Sarcastic, and the four 'Mixed type'—according to scores on the

TABLE 1

Humor Indices for Women

Item	Criterion		
	Kind of humor (b)	Amount of humor by the Wit (d)	
	(Per cent agreement)	Funny	Unfunny
Protocol group:			
One	71	2.15	1.44
Two	70	1.98	1.22
Wit type:			
Clowning	81	2.02	1.20
Sarcastic	80	2.03	1.12
Mixed*	51	2.14	1.67

* For (b), percentage agreement with E's type classification.

second index (b). The four in each type group thus derived were then subdivided into the two 'Funny' and the two 'Unfunny'—according to scores on the fourth index (d), adjusted for length of protocol. In Table 1 are presented the humor analysis scores for these groupings on the two relevant criteria—(b) and (d). Women subjects achieved 92 per cent agreement on Wit identification with practically no variation in score; scores on amount of humor paralleled those on (d) with an average of 2.26. The concern here was not one of statistically significant differences; the distance along the Funniness dimension between Funny and Unfunny protocols was simply maximized. Differences between the two Protocol Groups, consequent upon the random division, must be taken into account in the interpretation of the results of the experiment.

STATUS JUDGMENTS

In standard classroom situations each of 100 men and 50 women received a booklet of 12 protocols arranged in a constant random order (subject to the

restriction for Variation I that each protocol group of 6 contain 3 of each humor type). The subject was instructed to read each skit in the booklet and answer the following four questions about it by writing the name of the protocol actor of his choice in the indicated place on a separate answer sheet:

1. If I were to associate with these three people, which one would probably have the most power to influence me?
2. If I were to associate with these three people, which one would I probably like best?
3. Of these three people which one probably has the most power to influence the group?
4. Of these three people which one is probably best liked by the group?

Questions 1 and 2 provided perceptions of the power and popularity dimensions in terms of the responding subject himself (Self Referent); 3 and 4 in terms of the subject's estimate of the feelings of the protocol actors (Group Referent). In an exploratory study of this kind it seemed advisable to include both of these two obviously related but at least conceptually distinct tasks.

The material was presented to the subjects as a study in social perception—"how we form impressions of the people in a group and their feelings about each other." Humor was not mentioned as relevant or irrelevant to the research interest.

For Variation I with the 100 men (all of whom received the same 12 protocols), 25 subjects were randomly assigned to each of four Presentation Conditions. Descriptions, rationale, and discussion of the findings in respect to the Presentation Condition variables can be found in the original thesis (3). Any variance associated with Conditions has been systematically isolated in all statistical tests; furthermore, the equal number of subjects in each condition (on humor analysis as well as status judgments) effectively controlled for these effects over-all. The 50 women in Variation II were randomly assigned to the two different sets of 12 protocols, so that there were 25 subjects receiving each set.

RESULTS

The status judgment material was scored in terms of the number of choices of the Wit by a subject over all protocols of one humor type for Variation I, within humor type and Funniness level for Variation II. In Tables 2 and 3 are presented the relevant status judgment sums-of-choices for Variations I and II; in Table 4, a comparison of the choices of men with those of women on the twelve Variation I protocols.[6] If the Wit were not differentiated from his companions it would be expected that he would in each instance receive one third of the total choices made. The number of Wit choices expected by chance

[6] Comparison across subject-sex, it must be recalled, involves not simply the comparison of reactions of men to those of women but reactions also to stimulus situations in which the sex of the actors has changed.

TABLE 2

Number of Choices of the Wit by Men

Prediction	Clowning Wits	Sarcastic Wits
Would influence self	221	179
Would be liked by self	266	172
Does influence group	216	238
Is liked by group	276	180

would therefore be 200 for each entry in Table 2, 33 in Table 3, and 50 in Table 4.

An appropriate statistic for the status judgment data is analysis of variance —the design a balanced multiple-classification, mixed model. With choice of Wit per subject per question per class of protocols as the smallest subclass (terminology and notation is that of Henderson, [4]), and with independence assured for the only random variable, subjects, a sufficient number of observations was available to make deviation of the population distribution from normality of no practical consequence in either treatment.

Variation I

The results of the main analysis of the Variation I status judgments are presented in Table 5. For purposes of convenience and simplification, Presentation Condition was here treated as one variable; the terms involving it have been omitted from the table. The only random effect in the analysis is H, E—Subjects; the others are regarded as fixed (i.e., any inference restricted to the difference associated with the particular levels in the experiment). For each fixed effect in the analysis the appropriate denominator for the F ratio for testing the null hypothesis is the term expressing the interaction of that effect with the random effect. For example, the appropriate test of the hypothesis that $\sigma^2_{AB} = 0$ is $AB/ABH, E$.

TABLE 3

Number of Choices of the Wit by Women

Prediction	Clowning Wits		Sarcastic Wits		Mixed Wits	
	Funny	Unfunny	Funny	Unfunny	Funny	Unfunny
Would influence self	42	36	33	12	35	29
Would be liked by self	36	46	14	5	37	25
Does influence group	41	39	55	25	39	38
Is liked by group	40	40	20	9	31	31

TABLE 4

Comparison Men X Women on Number of Choices of the Wit for Twelve Variation I Protocols

Prediction	Clowning Wits		Sarcastic Wits	
	Men	Women	Men	Women
Would influence self	71	59	45	33
Would be liked by self	81	68	47	13
Does influence group	67	59	61	58
Is liked by group	81	64	44	20

NOTE: Scores are those for the 25 male subjects in Presentation Condition One (that which all women subjects received) and those contributed by 50 women subjects in all, 25 for any particular protocol.

Effects associated with two of the three main classifications were significant: type of humor (a Clowning Wit was chosen more than a Sarcastic Wit), and Referent (a Wit was chosen more when the referent was to intragroup perceptions than when it was to observer-self perceptions). The significant *AB* and *ABC* interactions refer directly to the main interest, the effects of being witty in a Clowning or a Sarcastic way upon position on the separate dimen-

TABLE 5

Analysis of Variance of Status Judgments by Men

Source of variation	df	Mean square	F	Significant level
A	1	55.125	27.494	.001
B	1	2.000	0.509	ns
C	1	6.480	3.983	.05
AB	1	36.125	24.948	.001
AC	1	4.805	3.884	ns
BC	1	1.620	1.176	ns
ABC	1	5.445	8.609	.01
H, E	96	8.491		
AH, E	96	2.005		
BH, E	96	3.932		
CH, E	96	1.627		
ABH, E	96	1.448		
ACH, E	96	1.237		
BCH, E	96	1.377		
ABCH, E	96	0.6325		

NOTE: A—Humor Type; B—Status Dimension; C—Referent; H, E—Subjects within Presentation Condition

sions of influence and popularity. When responses are considered separately for each humor type, it is clear that an excess of popularity over influence choices was indeed characteristic of the Clown, while for the Sarcastic Wit there was but slight differentiation between the two dimensions of status. When the additional factor of Referent (*ABC* interaction) is taken into account, however, the pattern for the Sarcastic Wit is clarified. As the sums in Table 2 show, the predicted excess of influence over popularity for the Sarcastic Wit did in fact occur, but only when the subject was reporting in terms of the interacting group. When the referent was to self, there was little difference between perceived influence and perceived popularity for the Sarcastic Wit.

A somewhat different question concerns the social position of the Wit as compared to his nonwitty companions; did being witty in the ways specified increase (or decrease) power or popularity beyond chance levels? For each of the entries in Table 2 the difference between observed and expected Wit-choice (in this case one third of the total choices, or 200) was tested by the t statistic.[7] As with the analysis of internal relative choice, results for these tests tended to support the hypotheses. The Clown was perceived as popular; his perceived influence was within chance limits. The Sarcastic Wit was not seen as popular, but he was definitely perceived as having influence over his group. Although not significant, the trend in the one remaining case—the reported influence of the Sarcastic Wit on the self—was in the opposite direction.

Variation II

This situation differed from Variation I in several respects in addition to the sex of the participants: an increase in the number of humor types from two to three; the appearance of a new variable, Funniness; and the substitution for the Presentation Condition variable (E) of the quite different Protocol Group classification (E′). Although the protocols were randomly divided into the two groups, the E′ effect is in this case more appropriately regarded as fixed (4, p. 16), leaving but one random variable (subjects). Again, each fixed effect is tested against its interaction with the random effect (H, E′). Results of this analysis for the four major factors and the interactions among them are presented in Table 6.

There are significant differences associated with each of the four major effects: type of humor (a Clown was chosen more than a Sarcastic Wit; choice of a Wit of Mixed type was intermediate); status dimension (a Wit was perceived as more influential than popular); referent (a Wit was chosen more when the referent was to intra-group perceptions than when it was to observer-

[7] Sums of squares for the variance estimates were computed about the four separate sample means (i.e., the four Presentation Conditions).

TABLE 6

Analysis of Variance of Status Judgments by Women

Source	df	Mean square	F	Significant
A	2	13.7908	22.086	.001
B	1	6.7500	6.459	.05
C	1	2.8034	5.902	.05
D	1	6.4534	10.960	.01
AB	2	4.4175	17.010	.001
AC	2	1.3408	6.048	.01
AD	2	3.5308	5.963	.01
BC	1	1.4700	5.852	.05
BD	1	1.6133	4.290	.05
CD	1	0.0000	0.000	ns
ABC	2	0.2775	1.023	ns
ABD	2	0.8308	2.541	ns
ACD	2	0.5576	2.153	ns
BCD	1	0.0000	0.000	ns
ABCD	2	0.3675	1.888	ns

Source	df	Mean square	Source	df	Mean square
H, E′	48	1.8894	BCH, E′	48	0.2512
AH, E′	96	0.6244	BDH, E′	48	0.3761
BH, E′	48	1.0451	CDH, E′	48	0.1490
CH, E′	48	0.4750	ABCH, E′	96	0.2712
DH, E′	48	0.5888	ABDH, E′	96	0.3269
ABH, E′	96	0.2597	ACDH, E′	96	0.2590
ACH, E′	96	0.2217	BCDH, E′	48	0.2225
ADH, E′	96	0.5921	ABCDH, E′	96	0.1946

NOTE: A—Humor Type; B—Status Dimension; C—Referent; D—Funniness; H, E′—Subjects within Protocol Group

self perceptions); and Funniness (a Funny Wit was chosen more than an Unfunny Wit). The five significant first-order interactions testify to the interrelated multivariate nature of Wit-choice. Differentiation between Funny and Unfunny Wits was greatest when the Wit's humor was Sarcastic (AD), and when the dimension of choice was influence (BD). Similarly the differentiation between Self and Group referent—greater choice on Group referent—was most pronounced for the Sarcastic Wit (AC) and for influence choices (BC). The highly significant interaction between type of humor and status dimension (AB) is of course the finding of most concern: while for the Clown there was little difference between perceived influence and perceived popularity, the Sarcastic Wit received many more choices as influential than as popular. Choice of the Wit of Mixed humor type remained consistently intermediate.

A discussion of the several significant effects associated with the Protocol Group (E′) factor is contained in (3); briefly, there was a fairly consistent underchoice by Group Two as compared to Group One in every instance. The complexity of the stimulus situations makes it seem highly likely that differences between the two groupings are attributable to differences among protocols (most noticeably the over-all difference in Funniness level apparent in Table 1) rather than to differences between the two groups of subjects.

The difference between observed and expected Wit-choice for each entry in Table 3 was tested (as with Variation I) by t, variance estimate in each case obtained by computing sums of squares about the two separate Protocol Group means. These results indicate that, for both the Clown and the Mixed Wit, choice deviated significantly from chance expectancy in only one of eight instances: choice of the Unfunny Wit on the self-referent popularity question. Choice of the Sarcastic Wit, on the other hand, was significantly different from chance in six of the eight instances. Whether funny or unfunny, this Wit type was definitely perceived as unpopular. The vital importance of Funniness level for the Sarcastic Wit shows up clearly in connection with the influence questions: as predicted, the Sarcastic Wit was reported as having influence over the group—but only if he was Funny. When the referent was to the observing subject himself, the perceived influence of the Funny Sarcastic Wit was not different from chance but the Unfunny Sarcastic Wit was perceived as significantly without influence.

Comparison by Sex

Examination of the status judgment choices of men and of women made under identical conditions and on the same group of protocols (Table 4) suggests that women consistently chose the Wit less than men did; that this underchoosing was pronounced when choice was for popularity, when a Sarcastic Wit was involved, and when the referent was to self. Chi square was used to test the eight separate sex-differences in Table 4. For only two of the cells was the difference by sex significant: in reference both to self and to group, the Sarcastic Wit was perceived as significantly less popular by the women than by the men.

DISCUSSION

At its most general level this study has demonstrated some relationships between 'being witty' irrespective of type of humor and 'being chosen' (or 'being visible' or 'being important') irrespective of dimension of choice. Since, however, the results have clearly established the importance of the type of humor employed by the Wit and the somewhat less incisive differentiation between the dimensions of social structure, it is clear that only when these

parameters are specified can findings be of much import (witness the essentially random results for Mixed humor type protocols in Variation II). Nevertheless one variable—Funniness—deserves consideration at this level. Tested in Variation II only, Funniness (amount of Wit-associated humor) was conceived of as essentially a validating device: the fact that changes in funniness were associated with changes in status is considered to be evidence for the contention that it was the 'wittiness' of the Wit which was the salient fact about him. That success in humor attempts was more crucial for a Sarcastic Wit is not surprising: unfunny whimsy may be charitably interpreted as indicative of good intentions; biting humor without the humor is plainly malice. It is probably the success-failure aspect of the funniness variable which is most responsible also for its greater impact over-all on the influence dimension.

When the selective effects of the methodological and the validating variables are taken into account, the general predictions that Clowning (whimsical, simple joking) is associated with being well liked and—if not with an actual lack—at least with an absence of any appreciable influence (power), and that being Sarcastic (biting, pointed ridicule) is on the other hand associated with being unusually influential (powerful) and—if not actively disliked—at least not well liked were substantially upheld by the data. Moreover, in spite of the double shift in sex of both responding subject and protocol actors, the same trends appeared for both sexes both in the humor analysis and in status judgments. It must of course be kept in mind that the findings refer throughout to interactional rather than to absolute differences: the relative position of one type of Wit on the two status dimensions or the relative position on one status dimension of the two types of Wit, and the position of the Wit in comparison to his non-Witty companions.

There were two major instances of discrepant results: in terms of the prediction there was a consistent underreporting of the influence of the Sarcastic Wit on the reporting self; and secondly, particularly on Variation II the Clown tended to be perceived as relatively influential. Explanations *post hoc* for these difficulties while easily devised would not seem particularly appropriate at this early stage in the understanding of the phenomena. It should rather be emphasized that the cited findings are valid only within the limits of the particular experimental setting. Generalization should not be extended beyond a college-student population, nor, strictly speaking, beyond the specific samples of an acting Wit occurring in the protocols here employed. In connection with this last stricture it is of some relevance that although there were in all but 24 different individual Wits, the number of separate occurrences of-humor (jokes) on which these results were based was considerably greater—attempts at humor by Wits totaled 216.

SUMMARY

The present study was concerned with one aspect of social humor—the effects of being witty upon the relationship between the Wit and his nonwitty associates. Hypotheses were stated as expected associations between the exercise of a particular type of humor and perceived position of the Wit on two dimensions of social structure. Specifically, it was predicted that a Sarcastic Wit would be perceived as powerful (influential) but unpopular, a Clowning Wit as popular but powerless (without influence).

Artificially simulated small group interactions in the form of written fictional conversation, each involving one person who consistently attempted to be humorous, were examined by 73 college student subjects who provided essentially validating measures of the degree of funniness of the Wit and the clarity of humor type in each of the situations. A series of status judgments, estimates of the position of the Wits on the two dimensions of social structure, were then made by 150 additional subjects.

The predicted relationships between being witty in the specified ways and position on the power and popularity dimensions were in general supported by the data. One methodological variable (whether judgments were made in terms of the judge or in terms of the judge's estimate of the feelings of the fictional group members) and one validating variable (degree of funniness of the Wit) exerted demonstrable and selective effects on the relationships—most noticeably a highly differential reaction to the influence dimension dependent on referent and a highly differential reaction to the Sarcastic Wit dependent on funniness level.

Manuscript received: December 2, 1958
Revised manuscript received: February 12, 1959
Jacqueline D. Goodchilds
Fels Group Dynamics Center
University of Delaware
Newark, Delaware

REFERENCES

1. Bales, R. F., *Interaction Process Analysis: A Method for the Study of Small Groups,* Cambridge, Mass.: Addison-Wesley, 1950.
2. Flugel, J. C., "Humor and Laughter," in G. Lindzey, ed., *Handbook of Social Psychology,* Vol. 2. *Special Fields and Applications,* Cambridge, Mass.: Addison-Wesley, 1954.
3. Goodchilds, J. D., "Effects of Being Witty on Position in the Social Structure of a Small Group," Ph.D. dissertation, Cornell University, 1957.
4. Henderson, C. R., "Analysis of Variance in Complex Balanced Multiple Classifications," (mimeographed paper) Cornell University, 1954.
5. Parsons, T., R. F. Bales, and E. A. Shils, *Working Papers in the Theory of Action,* Glencoe, Ill.: The Free Press, 1953.

Some Effects of Promotional Frustration on Employees' Understanding of, and Attitudes Toward, Management [1]

DAVID SIROTA, *The University of Michigan*

This study is designed to test an important assumption underlying some of the efforts, based on social psychological theory, now being undertaken in industry to improve organizational effectiveness. We refer to those aspects of the "human relations" approach which emphasize the importance of satisfying individual ego needs (e.g., needs for acceptance, approval, etc.) within the organization. For example, Likert (3), in his outline of a "modified theory of management," states:

> It is essential that . . . each interaction be of such a character that there will be a maximum probability that the subordinate will, in the light of his experience and expectations, feel that the experience is supportive and contributes to his sense of personal worth.

The theory postulates that the satisfaction of ego needs will result in attitudes of mutual confidence and trust. These will be conducive to full and frank communication between and within all organizational levels and should lead to a fuller understanding of the employees by management, and of management by the employees. According to this approach, frustration of the ego needs will therefore be detrimental to the operation of the communications network within an organization.

The present study is designed to provide a test of this prediction by investigating the relationship between the frustration of promotional aspirations (which, we assume, are strongly related to the ego needs) and communications effectiveness as measured by the amount of information employees have about management philosophy, goals, and operating procedures.

Since we are dealing with the kinds and amount of information employees have about management, it seems appropriate to conceptualize the problem in terms of cognitive processes. The essential question to be answered is: what are the effects of frustration on the cognitive field?

FRUSTRATION AND THE COGNITIVE FIELD

Krech and Crutchfield (2) postulate both "adaptive" and "maladaptive" consequences of frustration. The former category includes an active searching for new paths around the source of frustration—i.e., problem-solving responses.

[1] This study was conducted as part of a larger research project directed by D. C. Pelz, under the Organizational Behavior Program of the Survey Research Center. The writer wishes to thank Dr. Pelz for his generous assistance.

The complex of "maladaptive" responses is well known, e.g., regression, fixation, and aggression. These are maladaptive in the sense that they do not aid—in fact, they hinder—attainment of the original goal. The use of adaptive or maladaptive behaviors will depend partly on the severity of the frustration: the greater the severity the greater the probability of a maladaptive response.

In terms of the effects on the cognitive field, Krech and Crutchfield argue that adaptive behavior is accompanied by a widening of awareness as the individual seeks new information about alternative paths to the goal. The individual becomes sensitized to new features of the environment. The primary cognitive feature of maladaptive responses to frustration, on the other hand, is a narrowing of the field: "The person focuses his attention so completely on the blocked pathways for the inaccessible goal that he is blinded to the existence of alternative pathways or substitute goals." Maier (4), in his discussion of "frustration-instigated" behavior, arrives at approximately the same conclusion:

> The frustrated person . . . must be unusually aware of the mass of feelings or sensations of internal origin. When these sensations are excessive, they would occupy his attention and consequently make him less aware of the external world as well as insensitive to the feelings of other persons.

The above distinction may be applied to the analysis of the responses to promotional frustrations engendered within an industrial setting. Promotion ordinarily depends on the degree to which the employee conforms to the demands and expectations of management. His ability to do so will be conditioned by how much information he has regarding these demands. If there has been no promotional frustration, there will be little motivation to obtain such information. If promotional aspirations have been moderately frustrated, we would expect an "adaptive" response, i.e., an active search for this information. The cognitive components of a "maladaptive" response—which would emerge under conditions of severe frustration—would be characterized by a decreased knowledge of management expectations, i.e., a narrowing of awareness of the alternative paths around the promotional barrier.

Hypotheses

The above argument outlines the following curvilinear hypothesis:

Hypothesis 1. Moderately frustrated employees will have the greatest amount of information about management, satisfied employees somewhat less, while highly frustrated employees will have the least information.

What about the influence of frustration on attitudes toward the company and supervisor? The basic hypothesis in this area—repeatedly demonstrated—is that hostility is a direct function of frustration:

Hypothesis 2. The most highly frustrated employees will have the greatest amount of hostility toward management and the supervisor, the moderately frustrated less, and the satisfied the least.

The predictions deviate from Likert's in the following way: we predict that the amount of information about the company will not decrease directly with amount of hostility toward it; that some frustration, and its resulting hostility, might actually be conducive to increased awareness. However, too much frustration (and hostility) will have a "blinding" effect. On this latter issue Likert's predictions and the hypotheses presented here are in agreement.

PROCEDURE AND OPERATIONS

In September, 1957, a questionnaire was administered to 2201 employees in two divisions of an electronics manufacturing firm in a southern state.

We obtained measures of *employee information about management* in three areas. The first of these is *management philosophy*. Three sets of five statements each were presented to the employees, each statement representing a possible aspect of management philosophy; some statements did in fact represent actual philosophy, while others did not. Within each set the employees were asked to rank the statements as they thought the general management people would rank them. Examples are: "Jobs are secure with little danger of layoffs," "People at all levels are encouraged to express their opinions frankly without fear of 'stepping on toes,' " "[The company] wants to get into product lines requiring relatively little investment in facilities and equipment." Responses to the above statements were compared to a standard derived from an average of top management responses.

Information about *management goals* was confined to planned company expansion. Respondents were asked to estimate specific figures for company goals regarding sales and employment by the end of 1957. An accuracy score was derived by comparing the answers to the above two questions to the estimates made by top management.

We also obtained a measure of information about *personnel policies*. Eight true-false type questions were asked about current personnel policies. A score was computed based on the number of correct answers.

A number of questions were included relating to the employee's *attitude toward the company and his supervisor,* (e.g., "How does [the company] compare generally with other places as a place to work?"; "If you were offered another job [in the company] doing about the same type of work but under a different supervisor, how would you feel about moving?"). These questions were used to construct indices of attitude toward the company and attitude toward the supervisor.

The *promotional frustration* index was derived from answers to two ques-

TABLE 1

Promotional Frustration in Relation to Information about Management *

	N	Information about: (a) Management philosophy	(b) Personnel policies	(c) Expansion plans
Satisfied	1658	.37	5.50	2.41
Moderately frustrated	228	.41	5.90	2.49
Highly frustrated	310	.33	5.38	2.34
F:		5.45	9.46	5.13
P:		<.01	<.01	<.01

* Information about management philosophy (a) is represented by correlations between employee perceptions and management's answers; about personnel policies (b) by number correct in 8 true-false items; and about expansion plans (c) by a 3-point scale derived from similarities to management estimates.

tions: "Approximately how long do you think it will be before you are asked to take on a job at a higher grade level?" (*expected* time to next promotion) and "Approximately when would you like to take on a job at a higher grade level?" (*desired* time to next promotion). Both questions were followed by identical 6-point scales. The discrepancy between answers to the two questions was computed, and employees divided into three groups:

1. *Satisfied*—expected period is within 1 point of the desired period. (N=1658)
2. *Moderately frustrated*—desired period is greater than expected by 2 points. (N=228)
3. *Highly frustrated*—desired is greater than expected by more than 2 points. (N=310)

RESULTS AND CONCLUSIONS

Table 1 presents the average scores on the three dimensions of information about management for the three degrees of promotional frustration. Means for attitude toward the company and toward the supervisor are presented in Table 2.

Tables 1 and 2 indicate confirmation of the two hypotheses. The highly frustrated group tends to be lowest on information and highest on hostility; the moderately frustrated is highest on information and medium on hostility; the satisfied group exhibits the most positive attitudes toward the company but has only medium amounts of information.[2]

[2] Controls for job grade, age, education, and length of time employed by company do not significantly alter these results.

We have noted that Likert's approach emphasizes the necessity for fulfillment of the ego needs. There is little question that such fulfillment does result in a significant increase in satisfaction or "morale" (as our results indicate). However, the relationship between morale and organizational effectiveness has never been established; the research to date has yielded insignificant or contradictory findings (3). Our results suggest that a source of the difficulty lies in the oversimple assumption that job effectiveness is a direct function of need-fulfillment. In the area of promotional opportunities (and the ego needs which these represent) maximum effectiveness may accompany a moderate amount of frustration.

TABLE 2

Promotional Frustration in Relation to Attitude toward the Company and the Supervisor *

	N	Attitude toward: Company	Attitude toward: Supervisor
Satisfied	1658	4.12	4.01
Moderately frustrated	228	3.96	3.56
Highly frustrated	310	3.69	3.29
F:		46.41	33.30
P:		<.01	<.01

* Each attitude is measured by average response on three 5-point scales.

We would not, however, conclude that moderate amounts of frustration will always benefit the organization. The effects, obviously, will depend on the kinds of behaviors which the moderately frustrated individual is motivated to perform to achieve his goals. These behaviors may or may not increase the quality of his performance as a member of the organization. In the company investigated, moderate promotional frustration is accompanied by better understanding. If the goal is higher wages, however, the major "adaptive" kind of behavior may well be to strike—a highly disruptive response in terms of the goals of the organization. The general principle regarding the consequences of frustration for the individual can, therefore, be applied to organizational settings only after the compatibility between the elicited behaviors and the goals of the organization have been carefully spelled out.

Our view is similar to the "path-goal" formulation of Kahn and others (1):

> If a worker sees higher productivity as a path leading to the attainment of one or more of his personal goals, he will tend to be a high producer, provided that he is free to do so. Conversely, if he sees low productivity as a path to the achievement of his goals, he will tend to be a low producer.

The major qualification which we would add is that if goal-attainment is perceived to be extremely improbable, "maladaptive" rather than problem-solving behavior is likely to ensue. Kahn's formulation, as it stands, does not provide for irrational responses under conditions of high frustration.

SUMMARY

Questionnaires administered to employees in an industrial plant reveal that although promotional frustration is directly related to hostility toward the company and the supervisor, it is curvilinearly related to the amount of objective information about management philosophy, goals, and operating procedures: frustrated employees have the least information, moderately frustrated the most, while those who are satisfied have medium amounts. These findings are interpreted to cast doubt on the assumption often made in "human relations" theory that maximum fulfillment of ego needs results in the highest degree of organizational effectiveness.

Manuscript received: February 19, 1959

David Sirota
Survey Research Center
University of Michigan
Ann Arbor, Michigan

BIBLIOGRAPHY

1. Kahn, R. L., "Human Relations on the Shop Floor," in E. N. Hugh-Jones, ed., *Human Relations and Modern Management*, Amsterdam: North-Holland, 1958.
2. Krech, D., and R. S. Crutchfield, *Theory and Problems of Social Psychology*, New York: McGraw-Hill, 1948.
3. Likert, R., *Developing Patterns of Management: II*, General Management Series Number 182, American Management Association, 1956.
4. Maier, N. R. F., "Frustration Theory: Restatement and Extension," *Psychological Review*, 1956, 63, 370–388.